

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:26:20 ; Search time 10.8434 Seconds
(without alignments)
102.059 Million cell updates/sec

Title: US-09-230-111c-9

Perfect score: 26

Sequence: 1 TIQSVI 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 673684 seqs, 184443283 residues

Total number of hits satisfying chosen parameters: 673684

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	6	US-08-681-219-11	Sequence 11, Appl
2	26	100.0	6	US-10-092-138-9	Sequence 9, Appl
3	26	100.0	6	US-09-230-111c-9	Sequence 9, Appl
4	26	100.0	320	US-10-156-761-12666	Sequence 12666, A
5	25	96.2	322	US-09-815-242-5558	Sequence 5558, Ap
6	25	96.2	329	US-09-815-242-12506	Sequence 12506, A
7	25	96.2	329	US-10-301-997-24	Sequence 24, Appl
8	25	96.2	354	US-09-953-680A-19	Sequence 19, Appl
9	25	96.2	354	US-09-953-680A-21	Sequence 21, Appl
10	25	96.2	709	US-09-826-508-569	Sequence 589, App
11	24	92.3	103	US-10-187-394-11	Sequence 11, Appl
12	24	92.3	104	US-08-260-675-26	Sequence 26, Appl
13	24	92.3	104	US-10-050-050-26	Sequence 26, Appl
14	24	92.3	115	US-09-813-459-25	Sequence 25, Appl
15	24	92.3	115	US-09-813-459-25	Sequence 26, Appl

16	24	92.3	119	US-10-115-406-16	Sequence 16, Appl
17	24	92.3	119	US-10-154-333-18	Sequence 18, Appl
18	24	92.3	120	US-09-813-459-6	Sequence 6, Appl
19	24	92.3	120	US-09-813-459-16	Sequence 16, Appl
20	24	92.3	120	US-09-859-211-42	Sequence 42, Appl
21	24	92.3	120	US-09-880-708-20	Sequence 20, Appl
22	24	92.3	120	US-09-872-656-42	Sequence 42, Appl
23	24	92.3	120	US-10-335-483-24	Sequence 24, Appl
24	24	92.3	183	US-08-957-425-23	Sequence 23, Appl
25	24	92.3	472	US-10-005-228-7	Sequence 7, Appl
26	24	92.3	472	US-10-122-026-17	Sequence 17, Appl
27	24	92.3	473	US-09-813-398-25	Sequence 25, Appl
28	24	92.3	476	US-09-813-459-5	Sequence 5, Appl
29	24	92.3	479	US-09-813-398-26	Sequence 26, Appl
30	24	92.3	874	US-10-156-761-12986	Sequence 12986, A
31	24	92.3	914	US-10-128-714-8595	Sequence 8595, Ap
32	23	88.5	167	US-09-912-020-366	Sequence 366, App
33	23	88.5	198	US-09-848-813-2	Sequence 2, Appl
34	23	88.5	290	US-10-002-631C-32	Sequence 32, Appl
35	23	88.5	433	US-09-801-368-38	Sequence 38, Appl
36	23	88.5	549	US-09-726-643-61	Sequence 61, Appl
37	23	88.5	549	US-10-042-141-61	Sequence 61, Appl
38	23	88.5	609	US-09-866-050A-498	Sequence 498, App
39	23	88.5	610	US-09-989-723-113	Sequence 113, App
40	23	88.5	610	US-09-989-723-113	Sequence 113, App
41	23	88.5	610	US-09-989-727-113	Sequence 113, App
42	23	88.5	610	US-09-989-727-113	Sequence 113, App
43	23	88.5	610	US-09-989-731-113	Sequence 113, App
44	23	88.5	610	US-09-989-732-113	Sequence 113, App
45	23	88.5	610	US-09-991-073-113	Sequence 113, App

ALIGNMENTS

RESULT 1
US-08-681-219-11
; Sequence 11, Application US/08681219
; Publication No. US20020058607A1
; GENERAL INFORMATION:
; APPLICANT: Takaaki Sato and Junn Yanagisawa
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GPCR
; TITLE OF INVENTION: (PDE/DHR) DOMAIN AND USES THEREOF
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/681,219
; FILING DATE: 22-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/48962/JPW/JKM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-681-219-11

Query Match 100.0%; Score 26; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIGSVI 6
DB 1 TIGSVI 6

RESULT 2

US-10-092-138-9
Sequence 9; Application US/10092138
Publication No. US20030170723A1
GENERAL INFORMATION:
APPLICANT: Sato, Taka-Aki
TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
FILE REFERENCE: 65823/JPW/PT
CURRENT FILING DATE: 2002-09-06
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
US-10-092-138-9

Query Match 100.0%; Score 26; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIGSVI 6
DB 1 TIGSVI 6

RESULT 3

US-09-230-111c-9
Sequence 9; Application US/09230111C
Publication No. US20030203414A1
GENERAL INFORMATION:
APPLICANT: Sato, Taka-Aki
APPLICANT: Yamagisawa, Jun
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN
TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF (PDG/DHR)
FILE REFERENCE: 48962-A-PCF-US
CURRENT APPLICATION NUMBER: US/09/230,111C
CURRENT FILING DATE: 1999-05-17
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
US-09-230-111c-9

Query Match 100.0%; Score 26; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIGSVI 6

DB 1 TIGSVI 6

RESULT 4
US-10-156-761-12666
Sequence 12666; Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OKURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 12666
LENGTH: 320
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-12666

Query Match 100.0%; Score 26; DB 15; Length 320;
Best Local Similarity 100.0%; Pred. No. 1.18e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIGSVI 6
DB 244 TIGSVI 249

RESULT 5
US-09-815-242-5558
Sequence 5558; Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 5558
LENGTH: 322
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5558

Query Match 96.2%; Score 25; DB 9; Length 322;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIOSVI 6
126 TIQSVV 131

RESULT 6
US-09-815-242-12506
Sequence 12506, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zykand, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trewick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 12506
LENGTH: 329
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12506

Query Match 96.2%; Score 25; DB 9; Length 329;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIOSVI 6
128 TIQSVV 133

RESULT 7
US-10-301-997-24
Sequence 24, Application US/10301997
Publication No. US20030148346A1
GENERAL INFORMATION:
APPLICANT: Holden, David W.
TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS
NUMBER OF SEQUENCES: 106

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/301,997
FILING DATE: 22-Nov. US20030148346A1-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/527,431
FILING DATE: <Unknown>
APPLICATION NUMBER: US/08/887,534
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Rijn-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 28341/33996
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: (312) 474-6600

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 329 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 24:

US-10-301-997-24

Query Match 96.2%; Score 25; DB 12; Length 329;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIOSVI 6
128 TIQSVV 133

RESULT 8
US-09-952-680A-19
Sequence 19, Application US/09952680A
Publication No. US20030087239A1
GENERAL INFORMATION:

APPLICANT: Stanton, Marty
APPLICANT: Epstein, David
APPLICANT: Hamaguchi, No. US20030087239A1uko
TITLE OF INVENTION: Target Activated Biosensor and Methods of Using Same
FILE REFERENCE: 2339-501
CURRENT APPLICATION NUMBER: US/09/952,680A
CURRENT FILING DATE: 2001-09-13
PRIOR APPLICATION NUMBER: 60/232,454
PRIOR FILING DATE: 2000-09-13
NUMBER OF SEQ ID NOS: 75
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO: 19
LENGTH: 354
TYPE: PRT
ORGANISM: Homo sapiens
US-09-952-680A-19

Query Match 96.2%; Score 25; DB 11; Length 354;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TIOSVI 6
|||:
77 TIOSII 82

RESULT 9
US-09-952-680A-21

Sequence 21, Application US/09952680A
Publication No. US20030087239A1
GENERAL INFORMATION:
APPLICANT: Stanton, Marty
APPLICANT: Epstein, David
APPLICANT: Hamaguchi, No.
TITLE OF INVENTION: Target Activated Biosensor and Methods of Using Same
FILE REFERENCE: 23239-501
CURRENT APPLICATION NUMBER: US/09/952,680A
CURRENT FILING DATE: 2001-09-13
PRIOR APPLICATION NUMBER: 60/232,454
PRIOR FILING DATE: 2000-09-13
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21
LENGTH: 354
TYPE: PRT
ORGANISM: Homo sapiens
US-09-952-680A-21

Query Match 96.2%; Score 25; DB 11; Length 354;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TIOSVI 6
|||:
77 TIOSII 82

RESULT 10
US-09-826-509-589

Sequence 589, Application US/09826509
Publication No. US20030204073A1
GENERAL INFORMATION:
APPLICANT: Lehmann-Brunisma, Karin
APPLICANT: Liaw, Chen W.
APPLICANT: Lin, I-Lin
TITLE OF INVENTION: No. US20030204073A1-Endogenous, Constitutively Activated Known G
TITLE OF INVENTION: Protein-Coupled Receptors
FILE REFERENCE: AREN-207
CURRENT APPLICATION NUMBER: US/09/826,509
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/195,747
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
NUMBER OF SEQ ID NOS: 589
SOFTWARE: PatentIn Version 2.1
SEQ ID NO 589
LENGTH: 709
TYPE: PRT
ORGANISM: Homo sapiens
US-09-826-509-589

Query Match 96.2%; Score 25; DB 12; Length 709;
Best Local Similarity 83.3%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TIOSVI 6
|||:
432 TIOSII 437

RESULT 11
US-10-187-394-11
Sequence 11, Application US/10187394

Publication No. US20030176667A1

GENERAL INFORMATION:
APPLICANT: KRICK, PETER
APPLICANT: SMART, JOHN
TITLE OF INVENTION: SINGLE CHAIN ANALOGS OF TGF-B
TITLE OF INVENTION: SUPERFAMILY (MORPHONS)
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESS: PATENT ADMINISTRATOR, TESTA, HURWITZ &
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/187,394
FILING DATE: 28-JUNE-2002
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/496,398
FILING DATE: 02-FEB-2000
CLASSIFICATION:
APPLICATION NUMBER: US 08/478,097
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESQ, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: STK-059CN
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-248-7000
TELEFAX: 617-248-7100
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..103
OTHER INFORMATION: /note="BMP3 SEQUENCE"
US-10-187-394-11

Query Match 92.3%; Score 24; DB 12; Length 103;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TIOSVI 6
|||:
49 TIOSIV 54

RESULT 12
US-08-260-675-26

Sequence 26, Application US/08260675
Publication No. US2003010493A1
GENERAL INFORMATION:
APPLICANT: RUEGER, DAVID C
APPLICANT: KUBERSAMPATH, THANGAVEL
APPLICANT: OPPERMANN, HERMANN
APPLICANT: OZKAYNAK, ENGİN
APPLICANT: PANG, ROY HL
APPLICANT: COHEN, CHARLES M
TITLE OF INVENTION: MORPHOGEN-INDUCED NERVE REGENERATION AND
TITLE OF INVENTION: REPAIR

NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA HURWITZ & THIBEAULT
STREET: 55 STATE STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/260,675
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/126,100
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/922,813
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/667,274
FILING DATE: 11-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/752,764
FILING DATE: 30-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESQ. EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-070
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/248-7000
TELEFAX: 617/248-7100
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 104 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..104
OTHER INFORMATION: /note= "BMP3"
US-08-260-675-26

Query Match 92.3%; Score 24; DB 8; Length 104;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 TIOSVI 6
|||:
Db 49 TIOSIV 54

RESULT 13
US-10-050-050-26
Sequence 26, Application US/10050050
Publication NO. US20030125230A1
GENERAL INFORMATION:
APPLICANT: COHEN, CHARLES M.
CHARETTE, MARC F.
KUBERASAMPATH, THANGAVEL
RUEGER, DAVID C.
OPPERMANN, HERMANN
PANG, ROY H. L.
OZKAYNAK, ENGİN
SMART, JOHN E.
TITLE OF INVENTION: MORPHOGEN TREATMENT FOR LIMITING
PROLIFERATION OF EPITHELIAL CELLS.

NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
STREET: 45 SOUTH STREET
CITY: HOPKINTON
STATE: MA
COUNTRY: USA
ZIP: 01748
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/050,050
FILING DATE: 15-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/461,113
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/445,882
FILING DATE: 22-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FENTON Esq., GILLIAN M.
REGISTRATION NUMBER: 36,508
REFERENCE/DOCKET NUMBER: CRP-074DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 435-9001
TELEFAX: (508) 435-6951
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 104 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..104
OTHER INFORMATION: /label= BMP3
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-10-050-050-26

Query Match 92.3%; Score 24; DB 15; Length 104;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 TIOSVI 6
|||:
Db 49 TIOSIV 54

RESULT 14
US-09-813-459-25
Sequence 25, Application US/09813459
Patent NO. US20020107369A1
GENERAL INFORMATION:
APPLICANT: Lee, Se-jin
Cunningham, No. US20020107369A1
TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-10
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 1880 Century Park East, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/813,459
FILING DATE: 20-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/624,635
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr., Ph.D., John R.,
REGISTRATION NUMBER: 31,678
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 115 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
IMMEDIATE SOURCE: HUMAN GDF-10
CLONE: Human GDF-10
FEATURE:
NAME/KEY: Protein
LOCATION: 1..115
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-813-459-25

Query Match 92.3%; Score 24; DB 10; Length 115;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TIOSVI 6
Db 61 TIOSIV 66

RESULT 15
US-09-813-459-26
Sequence 26, Application US/09813459
Patent No. US20020107369A1
GENERAL INFORMATION:
APPLICANT: Lee, Se-jin
CUMMINGHAM, NO. US20020107369A1
TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-10
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Judas & Lubitz
STREET: 1880 Century Park East, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/813,459
FILING DATE: 20-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/624,635
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr., Ph.D., John R.,
REGISTRATION NUMBER: 31,678
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 115 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
IMMEDIATE SOURCE: MURINE GDF-10
CLONE: Murine GDF-10
FEATURE:
NAME/KEY: Protein
LOCATION: 1..115
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-813-459-26

Query Match 92.3%; Score 24; DB 10; Length 115;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TIOSVI 6
Db 61 TIOSIV 66

Search completed: November 26, 2003, 12:38:38
Job time : 10.8434 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:21:29 ; Search time 16.6265 Seconds
(without alignments)
57.280 Million cell updates/sec

Title: US-09-230-111c-9
Perfect score: 26
Sequence: 1 TIQSVI 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	6	19	AAW50180
2	26	100.0	15	281	AA44929
3	25	96.2	89	21	AAAG00992
4	25	96.2	322	22	AAU34062
5	25	96.2	329	20	AAW97691
6	25	96.2	329	22	AAU36913
7	25	96.2	332	23	ABP40384
8	25	96.2	354	21	AAV94207
9	25	96.2	354	21	AAV85290

10	25	96.2	354	22	AA899064	Human G-protein al
11	25	96.2	354	22	ABB99066	Human G-protein al
12	25	96.2	354	23	ABB09273	G protein-coupled
13	25	96.2	709	22	ABB56396	G protein-coupled
14	25	96.2	725	22	AA899036	TSHR-Gs-alpha fusi
15	25	96.2	14	15	AA851649	Human somatostatin
16	24	92.3	14	20	AAV43102	Osteogenic unglyco
17	24	92.3	14	22	AA850464	Bovine osteogenic
18	24	92.3	19	17	AAW03217	Bovine BMP-2 tryp
19	24	92.3	32	18	AAW24728	BMP-3 fragment 80-
20	24	92.3	63	9	AA81520	Heel domain of BMP
21	24	92.3	66	18	AAW30345	Bovine Bone Morpho
22	24	92.3	96	20	AAV48452	Fragment of bone m
23	24	92.3	103	10	AA891247	Human prostate can
24	24	92.3	103	10	AA895685	Protein CBP-3. M
25	24	92.3	103	15	AA847294	CBP-3 protein for
26	24	92.3	103	20	AAV16711	BMP3. Homo sapien
27	24	92.3	103	21	AA809529	WO9914235 Seq ID N
28	24	92.3	103	21	AA809529	Human BMP-3, SEO I
29	24	92.3	103	21	AA802795	Human BMP3 amino a
30	24	92.3	103	21	AAV92564	BMP 3 finger-1-hee
31	24	92.3	104	14	AA833414	Human BMP3 (fx) sev
32	24	92.3	104	14	AA833936	Morphogen BMP3. H
33	24	92.3	104	15	AA846746	Human BMP3 (fx) mo
34	24	92.3	104	15	AA857982	Human BMP3 (fx) H
35	24	92.3	104	15	AA850209	Human BMP3 (fx) H
36	24	92.3	104	15	AA860977	Human BMP3 (fx) H
37	24	92.3	104	15	AA860963	Human BMP3 (fx) H
38	24	92.3	104	17	AAW00233	Human BMP3 (fx) mor
39	24	92.3	104	18	AAW40196	Human BMP3 (fx) mor
40	24	92.3	104	18	AAW36865	Conserved 7 cystei
41	24	92.3	104	19	AAW48924	Human bone morphog
42	24	92.3	104	21	AAV57234	Human bone morphog
43	24	92.3	110	23	AAW51921	Human TGFbeta prot
44	24	92.3	110	23	AAW51938	Human TGFbeta prot
45	24	92.3	111	20	AAV08293	Human growth facto

ALIGNMENTS

RESULT 1
ID AAW50180 standard, peptide; 6 AA.
AAW50180;
AC AAW50180;
XX
XX
XX
DT 16-JUL-1998 (first entry)
XX
XX
DE Signal-transducing protein carboxy-terminal peptide.
XX
XX Inhibition; specific binding; signal-transducing protein;
XX cytoplasmic protein; proliferation; cancer cell; apoptosis;
KW virally infected cell.
XX
XX
OS Synthetic.
XX
XX WO9805347-A1.
XX
XX
PD 12-FEB-1998.
XX
XX
PF 18-JUL-1997; 97WO-US12677.
XX
XX
PR 22-JUL-1996; 96US-0681219.
XX
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
XX Sato T, Yanagisawa J;
XX
XX WPI; 1998-145347/13.
DR
XX
XX Inhibition of signal transduction - by inhibiting binding between a
PT signal-transducing protein and a cytoplasmic protein, for treating

PT e.g. cancer or viral infection
 XX
 PS Claim 22; Page 64; 108pp; English.
 CC A novel composition is capable of inhibiting specific binding
 CC between a signal-transducing protein (STP) having the
 CC carboxy-terminal sequence (Ser/Thr)-Xaa-(Val/Ile/Leu), where Xaa =
 CC any amino acid (e.g. the present peptide), and a cytoplasmic
 CC protein (CP) containing the sequence AAM50162 or AAM50163.
 CC The composition can be used to inhibit the proliferation of cancer
 CC or virally infected cells, or induce apoptosis in cancer or virally
 CC infected cells.
 CC
 SQ Sequence 6 AA;
 QY
 DB 1 TIOSVI 6
 1 TIOSVI 6
 1 TIOSVI 6
 RESULT 2
 AAR44929
 ID AAR44929 standard; Protein; 15281 AA.
 AC AAR44929;
 DT 25-MAR-2003 (updated)
 DT 08-JUL-1994 (first entry)
 DE T. niveum Cyclosporin synthetase.
 XX Enzyme; cyclosporin; synthetase-like activity; Tolypocladium niveum;
 KM T. inflatum GAMS; biosynthesis; vector; cyclosporin synthetase.
 XX
 OS Tolypocladium niveum.
 XX
 PN EP578616-A2.
 PD 12-JAN-1994.
 PF 05-JUL-1993; 93EP-0810474.
 XX
 XX 09-JUL-1992; 92AT-0001403.
 PR 08-MAR-1993; 93AT-0000437.
 PR 29-APR-1993; 93CH-0001310.
 PR 04-MAY-1993; 93CH-0001375.
 XX
 PA (SANO) SANDOZ LTD.
 PA (SANO) SANDOZ PATENT GMBH.
 PA (SANO) SANDOZ-ERFINDUNGEN VERW GES MBH.
 PI Leitner E, Schneider E, Schoergerdorfer K, Weber G;
 DR WPI; 1994-010432/02.
 DR N-PSDB; AAO54386.
 PT Isolated DNA sequence - which codes for enzyme having cyclosporin
 PT synthetase like activity
 XX
 PS Claim 1; Page 41-84; 93pp; English.
 XX
 CC This sequence represents an enzyme which has cyclosporin synthetase-
 CC like activity. This sequence was isolated from Tolypocladium niveum
 CC (formerly known as T. inflatum GAMS). This enzyme catalyses the
 CC peptide biosynthesis of cyclosporins and structurally related
 CC molecules. This sequence may be used for the production of
 CC cyclosporin by transforming a vector containing this sequence in
 CC to a recombinant host. This allows effective production of anti-
 CC biotic cyclosporin or its derivatives.

CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 15281 AA;
 QY
 DB 1 TIOSVI 6
 1 TIOSVI 6
 1 TIOSVI 6
 RESULT 3
 AAG00992
 ID AAG00992 standard; Protein; 89 AA.
 AC AAG00992;
 DT 06-OCT-2000 (first entry)
 DE Human secreted protein, SEQ ID NO: 5073.
 XX
 XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KM gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 PD 06-SEP-2000.
 PF 21-FEB-2000; 2000EP-0200610.
 PR 26-FEB-1999; 99US-0122487.
 PR (GEST) GENSET.
 XX
 PI Dumas Malne Edwards J, Duclert A, Giordano J;
 DR WPI; 2000-500381/45.
 DR N-PSDB; AAC00998.
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 PS Claim 13; SEQ ID 5073; 71pp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or poly(A) RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 CC
 SQ Sequence 89 AA;
 QY
 DB 1 TIOSVI 6
 1 TIOSVI 6
 1 TIOSVI 6
 Query Match 96.2%; Score 25; DB 21; Length 89;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 4
ID AAU34062 Standard; Protein; 322 AA.
XX
XX AAU34062;
AC
XX
XX
DT 14-FEB-2002 (first entry)
XX
XX Staphylococcus aureus cellular proliferation protein #338.
DE
XX
XX Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
XX Staphylococcus aureus.
OS
XX
XX WO200170955-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 21-MAR-2001; 2001WO-US09180.
PF
XX
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX
XX Hasebeck R, Ohlsen KL, Zykkind JW, Wall D, Trawick JD, Carr GJ,
PI Yamamoto RT, Xu HH;
XX
XX MPI; 2001-611495/70.
DR N-PSDB; AAG51921.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
PS
XX Example 3; Seq ID No 5558; 511pp; English.
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 322 AA;
SQ

Query Match 96.2%; Score 25; DB 22; Length 322;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
DB 126 TIQSVI 131

RESULT 5
ID AAW97691 Standard; Protein; 329 AA.
XX
XX AAW97691;
AC
XX
XX
DT 21-MAY-1999 (first entry)
XX
XX Staphylococcus aureus mutant P14C15 virulence gene product.
DE
XX
XX Virulence; P14C15; vaccine; antibacterial; antibiotic; screening;
KW bacteraemia; aspartate semialdehyde dehydrogenase.
XX
XX Staphylococcus aureus.
OS
XX
XX WO9901473-A2.
PN
XX
XX 14-JAN-1999.
PD
XX
XX 03-JUL-1998; 98WO-GB01974.
PF
XX
XX 03-JUL-1997; 97US-0887534.
PR
XX
XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
PA
XX
XX Holden DW;
PI
XX
XX MPI; 1999-105999/09.
DR N-PSDB; AAX07099.
XX
XX
PT Inhibition of virulence genes from *Staphylococcus aureus* - useful
PT for, e.g. screening for potential anti-microbial agents
PS Claim 1; Page 79-80; 203pp; English.
XX
XX This is the deduced amino acid sequence of the protein product of a
CC virulence gene (see AAX07099) newly identified in *Staphylococcus aureus*
CC mutant P14C15. *S. aureus* mutants containing chromosomal insertions
CC of a signature-tagged transposon were generated, and those with
CC attenuated virulence were identified in a mouse model of bacteraemia.
CC The nucleotide sequences of regions flanking the transposon insertion
CC sites of these mutants were determined, and database sequence
CC comparisons were performed to identify these virulence genes (see
CC AAX07088-116) and the possible function of their protein products
CC (see AAW97680-724). The P10C15 virulence gene product comprises the
CC N-terminus of aspartate semialdehyde dehydrogenase; a first ORF of
CC P10C15 encodes the C-terminus of the aspartokinase 2 alpha subunit
CC (see AAW97690). Both enzymes are involved in homoserine biosynthesis.
CC A claimed method of identifying an antibacterial agent involves
CC assaying potential agents for the ability to interfere with the
CC expression of *S. aureus* virulence gene products. Also new is a *S.*
CC *aureus* organism containing a functional mutation in one of the
CC virulence genes, and its use in vaccine compositions.
XX
XX Sequence 329 AA;
SQ

Query Match 96.2%; Score 25; DB 20; Length 329;
Best Local Similarity 83.3%; Pred. No. 4.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
DB 128 TIQSVI 133

RESULT 6
ID AAW36913 Standard; Protein; 329 AA.
XX
XX AAW36913;
AC
XX
XX
DT 14-FEB-2002 (first entry)

```
XX DE Staphylococcus aureus cellular proliferation protein #1093.
XX XX
XX KM Antisense; prokaryotic cellular proliferation protein;
XX KW antibiotic; antibacterial; drug design.
XX OS Staphylococcus aureus.
XX PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US09180.
XX PR 21-MAR-2000; 2000US-191078P.
XX PR 23-MAY-2000; 2000US-206848P.
XX PR 26-MAY-2000; 2000US-207727P.
XX PR 23-OCT-2000; 2000US-242578P.
XX PR 27-NOV-2000; 2000US-253625P.
XX PR 22-DEC-2000; 2000US-257931P.
XX PR 16-FEB-2001; 2001US-269308P.
XX PA (ELIT- ) ELITRA PHARM INC.
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ,
XX PI Yamamoto RT, Xu HH;
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS54772.
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids -
XX PS Example 3; Seq ID No 12506; 511pp; English.
XX XX
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the
XX CC genes, their use in the discovery of novel antibiotics, the essential
XX CC genes themselves and the encoded proteins. The prokaryotes used are
XX CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX CC invention is also useful for the identification of potential new targets
XX CC for antibiotic development. The antisense nucleic acids can also be used
XX CC to identify proteins used in proliferation, to express these proteins,
XX CC and to obtain antibodies capable of binding to the expressed proteins.
XX CC The proteins can be used to screen compounds in rational drug discovery
XX CC programmes. The antisense nucleic acid sequence is also useful to screen
XX CC for homologous nucleic acids which are required for cell proliferation in
XX CC a wide variety of organisms. The present sequence represents an
XX CC essential prokaryotic cellular proliferation protein.
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 329 AA;
XX
XX Query Match 96.2%; Score 25; DB 22; Length 329;
XX Best Local Similarity 83.3%; Pred. No. 4.7e+02;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
XX DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5229.
XX XX
XX KM Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
XX KW antibacterial; gene therapy.
XX OS Staphylococcus epidermidis.
XX PN US6380370-B1.
XX PD 30-APR-2002.
XX PF 13-AUG-1998; 98US-0134001.
XX PR 14-AUG-1997; 97US-055779P.
XX PR 08-NOV-1997; 97US-064964P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Doucette-Stamm LA, Bush D;
XX PI WPI; 2002-381255/41.
XX DR N-PSDB; ABN92929.
XX PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
XX PT polypeptide, useful for diagnosing and treating bacterial infections -
XX PS Disclosure; SEQ ID 5229; 267pp; English.
XX XX
XX CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
XX CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
XX CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
XX CC antibacterial activity and can be used in gene therapy. The sequences
XX CC can also be used in the diagnosis and treatment of bacterial infections,
XX CC particularly S. epidermidis infections. The sequences can be used to
XX CC screen for compounds able to interfere with the S. epidermidis life
XX CC cycle or inhibit S. epidermidis infection.
XX CC N.B. The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from the
XX CC USPTO web site.
XX SQ Sequence 332 AA;
XX
XX Query Match 96.2%; Score 25; DB 23; Length 332;
XX Best Local Similarity 83.3%; Pred. No. 4.8e+02;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

XX QY 1 TIQSVI 6
XX |||||:
XX 131 TIQSVV 136

XX DB 131 TIQSVV 136

XX RESULT 8
XX ID AAY94207 standard; Protein; 354 AA.
XX AC AAY94207;
XX XX
XX DT 31-JUL-2000 (first entry)
XX XX
XX DE Human G-alpha-13.
XX XX
XX KM Human; G-alpha-13; G protein; Gi protein; adenylyl cyclase;
XX KW dopamine; thyrotropin-releasing hormone; somatostatin;
XX KW signal transduction pathway.
XX XX
XX OS Homo sapiens.
XX PN US6063626-A.
XX PD 16-MAY-2000.
XX PF 24-JUN-1999; 99US-0339775.

[illegible]

G-protein coupled receptor (GPCR). This can be used in the treatment of

CC diseases, including hormone-dependent cancers, cardiovascular, nervous
 CC system, digestive system, immune system, respiratory, skeletal,
 CC endocrine, sensory and muscle diseases and disorders. The present
 CC sequence is a protein described in the exemplification of the invention.
 XX
 SQ Sequence 354 AA;

Query Match Score 25; DB 22; Length 354;
 Best Local Similarity 83.3%; Pred. No. 5.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIGSVI 6
 |||||
 Db 77 TIQSTII 82

RESULT 11
 AAB99066
 ID AAB99066 standard; Protein; 354 AA.

XX AAB99066;
 AC
 DT 23-AUG-2001 (first entry)
 XX
 DE Human G-protein alpha subunit 13.
 XX

KM G-protein coupled receptor; GPCR; GNR1 receptor; disease treatment;
 KM gonadotropin releasing; hormone receptor; hormone dependent cancer;
 KM human; catfish; goldfish; cow; sheep; horse; fruitfly; pig; rat; mouse;
 KM gene therapy.
 XX

OS Homo sapiens.
 XX
 FN WO200136446-A2.
 XX

PD 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000MO-GB04385.
 XX

PR 17-NOV-1999; 99GB-0027215.
 XX
 PA (UYBR-) UNIV BRISTOL.
 XX

PI McCardle CA;
 XX
 DR WPI; 2001-355607/37.
 XX

PT Use of a vector encoding G-protein coupled receptors for manufacturing
 PT medicaments for treating cancer, diseases of cardiovascular system,
 PT nervous system, digestive system, immune system, or muscle diseases -
 XX
 PS Disclosure; Fig 19; 78pp; English.
 XX

CC The present invention describes a produg comprising a vector encoding a
 CC G-protein coupled receptor (GPCR). This can be used in the treatment of
 CC diseases, including hormone-dependent cancers, cardiovascular, nervous
 CC system, digestive system, immune system, respiratory, skeletal,
 CC endocrine, sensory and muscle diseases and disorders. The present
 CC sequence is a protein described in the exemplification of the invention.
 CC
 XX

SQ Sequence 354 AA;
 Query Match Score 25; DB 22; Length 354;
 Best Local Similarity 83.3%; Pred. No. 5.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIGSVI 6
 |||||
 Db 77 TIQSTII 82

RESULT 12
 AAB09273

ID AAB09273 standard; Protein; 354 AA.

XX AAB09273;

DT 10-JUL-2002 (first entry)

DE G protein-coupled receptor (GPCR) 11 SEQ ID NO:19.

KM Target activated nucleic acid biosensor; signalling moiety; GPCR;
 KM nucleic acid sensor; detection; engineering; drug optimisation;
 KM G protein-coupled receptor.
 XX

OS Homo sapiens.
 XX

PN WO200222882-A2.
 XX

PD 21-MAR-2002.
 XX

PF 13-SEP-2001; 2001WO-US28835.
 XX

PR 13-SEP-2000; 2000US-232454P.
 XX

PA (ARCH-) ARCHEMIX CORP.
 XX

PI Stanton M, Epstein D, Hamaguchi N;
 XX

DR WPI; 2002-393977/42.
 XX

PT Nucleic acid sensor for detecting target molecule, comprises target
 PT molecule activation site and optical signalling unit that changes its
 PT optical properties upon allosteric modulation sensor after recognition
 PT of target -
 XX

PS Example 12; Page 89; 144pp; English.
 XX

CC The present invention describes a nucleic acid sensor molecule (I)
 CC comprising a target molecule activation site comprising a structure that
 CC recognises a target molecule and an optical signalling unit including at
 CC least one nucleotide coupled to a signalling moiety that changes its
 CC optical properties upon allosteric modulation of (I) following
 CC recognition of the target molecule. (I) is useful for detecting a target
 CC molecule associated with a pathological condition or genetic alteration.
 CC (I) is useful for identifying a drug compound, by identifying a nucleic
 CC acid biosensor-based molecule profile of target molecules associated with
 CC a disease trait in a patient, administering a candidate compound to the
 CC patient, and monitoring changes in the profile. Alternately, the method
 CC involves identifying a number of pathway target molecules, administering
 CC a candidate compound to a patient having a disease trait, and monitoring
 CC changes in the structure, level or activity of two or more of the pathway
 CC target molecules using (I). The profile of target molecules or the
 CC changes in the structure is compared to the profile of a reference
 CC healthy or diseased population. (I) is useful in multiple assays, for the
 CC detection of target molecule. (I) is also useful in diagnostic
 CC applications and drug optimisation. The present sequence represents a
 CC G protein-coupled receptor, which is used in an example from the present
 CC invention.
 XX

SQ Sequence 354 AA;
 Query Match Score 25; DB 23; Length 354;
 Best Local Similarity 83.3%; Pred. No. 5.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIGSVI 6
 |||||
 Db 77 TIQSTII 82

RESULT 13
 AAB09275
 ID AAB09275 standard; Protein; 354 AA.
 XX AAB09275;
 AC


```

XX 17-NOV-1999; 99GB-0027215.
PR (UVRB-) UNIV BRISTOL.
XX
PA Mcardle CA;
XX
PI WPI; 2001-355607/37.
XX
DR Use of a vector encoding G-protein coupled receptors for manufacturing
XX medicaments for treating cancer, diseases of cardiovascular system,
XX nervous system, digestive system, immune system, or muscle diseases
PT
PS Disclosure; Page 30-31; 78pp; English.
XX
CC The present invention describes a prodruq comprising a vector encoding a
CC G-protein coupled receptor (GPCR). This can be used in the treatment of
CC diseases, including hormone-dependent cancers, cardiovascular, nervous
CC system, digestive system, immune system, respiratory, skeletal,
CC endocrine, sensory and muscle diseases and disorders. The present
CC sequence is a protein described in the exemplification of the invention.
XX
SQ Sequence 725 AA;

```

Query Match 96.2%; Score 25; DB 22; Length 725;
 Best Local Similarity 83.3%; Pred. No. 1.1e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TIOSVI 6
   ||||:|
Db 448 TIOSII 453

```

Search completed: November 26, 2003, 12:30:20
 Job time : 17.6265 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:22:05 ; Search time 2.89157 Seconds

(without alignments)
97.580 Million cell updates/sec

Title: US-09-230-111c-9

Perfect score: 26

Sequence: 1 TIQSVI 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	96.2	304	1 YDEQ_ECOLI	P77588 escherichia
2	25	96.2	345	1 GBAK_XENLA	P27045 xenopus lae
3	25	96.2	353	1 GBAK_CAVPO	P38403 cavia porce
4	25	96.2	353	1 GBAK_HUMAN	P08754 homo sapien
5	25	96.2	353	1 GBAK_MOUSE	Q94511 mus musculu
6	25	96.2	353	1 GBAK_RAT	P08753 rattus norv
7	25	96.2	353	1 GBAK_CAVPO	P38401 cavia porce
8	25	96.2	353	1 GBAK_CHICK	P50146 gallus gall
9	25	96.2	353	1 GBAK_HUMAN	P04898 homo sapien
10	25	96.2	353	1 GBAK_ORYLA	P87383 oryzae lat
11	25	96.2	353	1 GBAK_RAT	P10824 rattus norv
12	25	96.2	353	1 GBAK_XENLA	P27044 xenopus lae
13	25	96.2	410	1 PRTP_HSV62	P52544 human herpe
14	25	96.2	504	1 VPS_AHSV6	O71026 african hor
15	25	96.2	505	1 VPS_AHSV9	O96597 african hor
16	25	96.2	872	1 SVY_STRCO	O06851 streptomyce
17	25	96.2	1522	1 DNA2_YEAST	P38859 saccharomyc
18	25	96.2	2303	1 POLG_TMEVB	P08844 t genome po
19	25	96.2	301	1 CDA2_YEAST	Q06702 saccharomyc
20	25	96.2	303	1 ACP7_MOUSE	O54794 mus musculu
21	25	96.2	309	1 DDL_BUCBP	P59435 buchnera ap
22	25	96.2	468	1 BMP3_RAT	P49002 rattus norv
23	25	96.2	472	1 BMP3_HUMAN	P12845 homo sapien
24	25	96.2	476	1 BMP3_MOUSE	P97737 mus musculu
25	25	96.2	476	1 BMP3_HUMAN	P55107 homo sapien
26	25	96.2	476	1 BMP3_HUMAN	P55108 rattus norv
27	25	96.2	476	1 BMP3_HUMAN	P55109 rattus norv
28	25	96.2	476	1 BMP3_HUMAN	P55110 rattus norv
29	25	96.2	476	1 BMP3_HUMAN	P55111 rattus norv
30	25	96.2	476	1 BMP3_HUMAN	P55112 rattus norv
31	25	96.2	476	1 BMP3_HUMAN	P55113 rattus norv
32	25	96.2	476	1 BMP3_HUMAN	P55114 rattus norv
33	25	96.2	476	1 BMP3_HUMAN	P55115 rattus norv

34	23	88.5	148	1 YF93_NEIMA	O51134 neisseria m
35	23	88.5	167	1 PFMG_ECOLI	P08190 escherichia
36	23	88.5	198	1 VIF_BIV06	P19562 bovine immu
37	23	88.5	198	1 VIF_BIV27	P19563 bovine immu
38	23	88.5	247	1 RHIR_RHILV	Q03316 rhizobium i
39	23	88.5	269	1 PPG_VIBVU	O08336 vibrio vuln
40	23	88.5	273	1 HMX2_MOUSE	P43687 mus musculu
41	23	88.5	353	1 GBI_ASTPE	P30676 aequorea pe
42	23	88.5	354	1 GBI2_MOUSE	P08752 mus musculu
43	23	88.5	356	1 MURB_BUCAI	P57153 buchnera ap
44	23	88.5	380	1 GRK_BACCH	Q92992 bacillus ha
45	23	88.5	417	1 MUA2_CLOPE	O8xh79 clostridium

ALIGNMENTS

RESULT 1
YDEQ_ECOLI
ID YDEQ_ECOLI STANDARD; PRT; 304 AA.
AC P77588;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical fimbrial-like protein ydeq precursor.
GN YDEQ OR B1502.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.,
RT "The complete genome sequence of Escherichia coli K-12";
RT Science 277:1453-1474 (1997).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=K12;
RC MEDLINE=97251357; PubMed=9097039;
RX Alde H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Samped G., Seki Y., Sivasubraman S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horituchi T.,
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377 (1996).
CC -1- SIMILARITY: TO TYPE-1 FIMBRIAL SUBUNITS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000247; AAC74575.1;
CC EMBL; D90792; BAA15175.1;
CC EMBL; D90793; BAA15183.1;
CC PIR; A64904; A64904.
CC Ecogen; EG13799; ydeQ.
CC InterPro; IPR000259; Fimbrial.
CC Pfam; PF00419; Fimbrial; 1.
CC Hypothetical protein; Fimbrial; Signal; Complete proteome.
CC SIGNAL 1 26
CC CHAIN 27 304
CC HYPOTHETICAL FIMBRIAL-LIKE PROTEIN YDEQ.

SQ SEQUENCE 304 AA; 32069 MW; 8153C86E3087D99A CRC64;

Query Match 96.2%; Score 25; DB 1; Length 304;

Best Local Similarity 83.3%; Pred. No. 33;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
|:|:|:|
DB 292 TIQSVI 297

RESULT 2

GBAK_XENLA STANDARD; PRT; 345 AA.

ID_GBAK_XENLA

AC P27045; 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DE 28-FEB-2003 (Rel. 41, Last annotation update)

GN Guanine nucleotide-binding protein G(k), alpha subunit (G(i) alpha-3)

OC (Fragment).

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI_TaxID=83355;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Oocyte;

RX MEDLINE=90346157; PubMed=2116977;

RA Olave J., Martinez S., Purcell P., Jorquera H., Codina J.,

RA Birbaumer L., Allende J.E.,

RT "Molecular cloning and sequence determination of four different cDNA

RT species coding for alpha-subunits of G proteins from Xenopus laevis

RT oocytes.";

RT FEBS Lett. 268:27-31(1990).

CC -1- FUNCTION: Guanine nucleotide-binding proteins (G proteins) are

CC involved as modulators or transducers in various transmembrane

CC signaling systems. G(k) is the stimulatory G protein of receptor-

CC regulated K(+) channels (By similarity).

CC -1- SUBUNIT: G proteins are composed of 3 units (alpha, beta and

CC gamma). The alpha chain contains the guanine nucleotide binding

CC site.

CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1

CC (G(i)/O/T(2)).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; X56090; CAA39570.1; -

CC DR PIR; S11046; RGXLI3.

CC DR HSSP; P10824; 1BOF.

CC DR InterPro; IPR001019; Gproteins_alpha.

CC DR Pfam; PF00503; G-alpha; 1.

CC DR PRINTS; PR00318; GPROTEIN.

CC DR ProDom; PD000281; Gproteins_alpha; 1.

CC DR SMART; SM00275; G_alpha; 1.

CC KW GTP-binding; Transducer; ADP-ribosylation; Multigene family.

CC FT NON_TER 1

CC FT NP_BIND 31 38 GTP (BY SIMILARITY).

CC FT NP_BIND 191 195 GTP (BY SIMILARITY).

CC FT NP_BIND 260 263 GTP (BY SIMILARITY).

CC FT MOD_RES 169 169 ADP-RIBOSYL[1] (BY ACTION OF CTX).

CC FT MOD_RES 342 342 ADP-RIBOSYL[1] (BY ACTION OF IAP).

CC SQ SEQUENCE 345 AA; 39648 MW; 74E070802EBA52 CRC64;

Query Match 96.2%; Score 25; DB 1; Length 345;

Best Local Similarity 83.3%; Pred. No. 38;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
|:|:|:|
DB 68 TIQSII 73

RESULT 3

GBAK_CAVPO STANDARD; PRT; 353 AA.

ID_GBAK_CAVPO

AC P38403; 01-OCT-1994 (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Guanine nucleotide-binding protein G(k), alpha subunit (G(i) alpha-3).

GN GNAI3.

OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Caviidae; Cavia.

OX NCBI_TaxID=10141;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Hartley; TISSUE=Lung;

RX MEDLINE=93129640; PubMed=1482697;

RA Sakanaka C., Izumi T., Nakamura M., Honda Z.-I., Watanabe T.,

RA Minami M., Mutoh H., Bito H., Seyama Y., Ui M., Shimizu T.,

RT "Three types of G protein of the guinea-pig lung: cDNA cloning

RT and analysis of their tissue distribution."

RT Biochim. Biophys. Acta 1175:61-66(1992).

CC -1- FUNCTION: Guanine nucleotide-binding proteins (G proteins) are

CC involved as modulators or transducers in various transmembrane

CC signaling systems. G(k) is the stimulatory G protein of receptor-

CC regulated K(+) channels.

CC -1- SUBUNIT: G proteins are composed of 3 units (alpha, beta and

CC gamma). The alpha chain contains the guanine nucleotide binding

CC site.

CC -1- TISSUE SPECIFICITY: Ubiquitously expressed.

CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1

CC (G(i)/O/T(2)).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; D21234; BAA04766.1; -

CC DR HSSP; P10824; 1A83.

CC DR InterPro; IPR001019; Gproteins_alpha.

CC DR Pfam; PF00503; G-alpha; 1.

CC DR PRINTS; PR00318; GPROTEIN.

CC DR ProDom; PD000281; Gproteins_alpha; 1.

CC DR SMART; SM00275; G_alpha; 1.

CC KW GTP-binding; Transducer; ADP-ribosylation; Multigene family;

CC KW Myristate.

CC FT INIT_MER 0 0 BY SIMILARITY.

CC FT LIPID 1 1 MYRISTATE (BY SIMILARITY).

CC FT NP_BIND 39 46 GTP (BY SIMILARITY).

CC FT NP_BIND 199 203 GTP (BY SIMILARITY).

CC FT NP_BIND 268 271 GTP (BY SIMILARITY).

CC FT MOD_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX).

CC FT MOD_RES 350 350 ADP-RIBOSYL[1] (BY ACTION OF PTX).

CC SQ SEQUENCE 353 AA; 40473 MW; E54FC805CEB97BBE CRC64;

Query Match 96.2%; Score 25; DB 1; Length 353;

Best Local Similarity 83.3%; Pred. No. 39;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
|:|:|:|
DB 76 TIQSII 81

ID	GBAK_HUMAN	STANDARD;	PRT;	353 AA.
AC	P08754	P17539/		
DT	01-NOV-1988	(Rel. 09, Created)		
DT	01-OCT-1994	(Rel. 30, Last sequence update)		
DT	15-SEP-2003	(Rel. 42, Last annotation update)		
DE	Guanine nucleotide-binding protein G(k), alpha subunit (G(i) alpha-3)			
GN	GNM13			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=86199230; Pubmed=2634384;			
RT	Itoh H., Toyama R., Kozasa T., Tsukamoto T., Matsunaka M., Kaziro Y.,			
RT	subunit. Structure of three distinct molecular species of G protein alpha			
RL	Proc. Natl. Acad. Sci. U.S.A. 84:7886-7890(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=86066503; Pubmed=3120178;			
RT	Beale C.R., Wilson C.B., Perlmutter R.M.;			
RT	Proc. Natl. Acad. Sci. U.S.A. 84:7886-7890(1987).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=86199244; Pubmed=2453165;			
RA	Codina U., Olate J., Abramowitz J., Mattera R., Cook R.G.,			
RT	Birnbaumer L.;			
RT	"Alpha 1-3 cDNA encodes the alpha subunit of Gk, the stimulatory G			
RL	protein of receptor-regulated K+ channels.";			
RN	J. Biol. Chem. 263:6746-6750(1988).			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=88247980; Pubmed=3132707;			
RA	Kim S., Ang S.L., Bloch D.B., Bloch K.D., Kawahara Y., Tolman C.,			
RT	Lee R., Seidman J.G., Neer E.J.;			
RT	"Identification of cDNA encoding an additional alpha subunit of a			
RT	human GTP-binding protein: expression of three alpha 1 subtypes in			
RL	human tissues and cell lines.";			
RN	Proc. Natl. Acad. Sci. U.S.A. 85:4153-4157(1988).			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=87247315; Pubmed=3109953;			
RA	Didbury J.R., Snyderman R.;			
RT	"Molecular cloning of a new human G protein. Evidence for two G			
RT	alpha-like protein families.";			
RL	FEBS Lett. 219:259-263(1987).			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RA	Puhl H.L. Iii, Ikeda S.R., Aronstam R.S.;			
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=82388257; Pubmed=12477932;			
RA	Strauberg R.L., Feinold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buecker K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,			
RA	Dattlenko L., Manuella K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Rachals S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA				

```

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallie D.E.,
RA Scherch A., Schein J.E., Jones S.J.W., Maira M.A.,
RA "Generation and initial analysis of more than 15,000 full-length
RA human and mouse cDNA sequences."
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Guanine nucleotide-binding proteins (G proteins) are
CC involved as modulators of transducers in various transmembrane
CC signaling systems. G(k) is the stimulatory G protein of receptor-
CC regulated K(+) channels.
CC -1- SUBUNIT: G proteins are composed of 3 units (alpha, beta and
CC gamma). The alpha chain contains the guanine nucleotide binding
CC site.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1
CC (G1/O/T/Z)).
CC -----
CC THE SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; M20604; AAA35895.1; -.
CC EMBL; M20597; AAA35895.1; JOINED.
CC EMBL; M20598; AAA35895.1; JOINED.
CC EMBL; M20599; AAA35895.1; JOINED.
CC EMBL; M20600; AAA35895.1; JOINED.
CC EMBL; M20601; AAA35895.1; JOINED.
CC EMBL; M20602; AAA35895.1; JOINED.
CC EMBL; M20603; AAA35895.1; JOINED.
CC EMBL; J03198; AAA35896.1; -.
CC EMBL; J03199; AAA35896.1; -.
CC EMBL; M27543; AAA35899.1; -.
CC EMBL; AF493907; AA012621.1; -.
CC EMBL; BC025285; AA025285.1; -.
CC PIR; S02348; RGHU13.
CC HSSP; P10824; IAS3.
CC Genew; HGNC:4387; GNAI3.
CC MIM; 139370; -.
CC DR GO; GO:000263; P:heterotrimeric G-protein GTPase, alpha-subunit; TAS.
CC DR GO; GO:0007194; P:negative regulation of adenylate cyclase ac. .; TAS.
CC DR InterPro; IPR001019; Gprotein_alpha.
CC DR Pfam; PF00503; G-alpha; 1.
CC DR PRINTS; PR00318; GPROTEIN.
CC DR ProDom; PD000281; Gprotein_alpha; 1.
CC DR SMART; SM00215; G_alpha; 1.
CC KW GTP-binding; Transducer; ADP-ribosylation; Multigene family;
CC MW Myristate.
CC FT INIT_MET 0 0 BY SIMILARITY.
CC FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
CC FT NP_BIND 39 46 GTP (BY SIMILARITY).
CC FT NP_BIND 199 203 GTP (BY SIMILARITY).
CC FT NP_BIND 268 271 GTP (BY SIMILARITY).
CC FT MOD_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX).
CC FT MOD_RES 350 350 ADP-RIBOSYL[1] (BY ACTION OF IAP).
CC SO SEQUENCE 353 AA; 40401 MW; C1684127F9F398BF CRC64;

Query Match 96.2%; Score 25; DB 1; Length 353;
Best Local Similarity 83.3%; Pred. No. 39;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TIGOSVI 6
DB 76 TIGOSII 81

RESULT 5
DBAK_MOUSE STANDARD; PRT; 353 AA.
ID _GBAK_MOUSE

```

AC 09DCS1; 061019;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Guanine nucleotide-binding protein G(K), alpha subunit (G(i) alpha-3).
 GN GNA13.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CF-1; TISSUE=Lung;
 RX MEDLINE=21085660; PubMed=11217951;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Komoto H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glasl C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Maehima J., Mazzarelli J., Membrants P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitlaker C., Wilming L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N;
 RX MEDLINE=2388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klauener R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Bueltow K.H., Scheefer C.F., Hsieh N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Narsisina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullan S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Maira M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN (3)
 RP SEQUENCE OF 247-321 FROM N.A.
 RC STRAIN=CF-1; Harlan;
 RX MEDLINE=97011591; PubMed=8858601;
 RA Williams C.J., Schultz R.M., Kopf G.S.;
 RT "G protein gene expression during mouse oocyte growth and maturation,
 and preimplantation embryo development.";
 RL Mol. Reprod. Dev. 44:315-323(1996).
 CC -1- FUNCTION: Guanine nucleotide-binding proteins (G proteins) are
 CC involved as modulators or transducers in various transmembrane
 CC signaling systems. G(K) is the stimulatory G protein of receptor-
 CC regulated K(+) channels (By similarity).
 CC -1- SUBUNIT: G proteins are composed of 3 units (alpha, beta and
 CC gamma). The alpha chain contains the guanine nucleotide binding
 CC site.

CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1
 CC (G(i)O/T/Z).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AK004566; BAB23377.1; -
 CC EMBL; AK077490; BAC36828.1; -
 CC EMBL; BC041107; AAH41107.1; -
 CC EMBL; U38502; AAB01733.1; -
 CC HSSP; P10824; IAS3.
 CC MGD; MG1:95773; Gna13.
 CC GO; GO:0005794; C:Golgi apparatus; IDA.
 CC InterPro; IPR001019; Gprotein_alpha.
 CC Pfam; PF00503; G-alpha; 1.
 CC PRINTS; PR00318; GPROTEIN.
 CC ProDom; PD000281; Gprotein_alpha; 1.
 CC SMART; SM00275; G_alpha; 1.
 CC GTP-binding; Transducer; ADP-ribosylation; Multigene family;
 CC Myristate; Palmitate; Lipoprotein.
 CC
 CC INIT MET 0 0
 CC BY SIMILARITY.
 CC MYRISTATE (BY SIMILARITY).
 CC PALMITATE (BY SIMILARITY).
 CC FT LIPID 1 1
 CC FT NP_BIND 2 2
 CC FT NP_BIND 39 46
 CC FT NP_BIND 199 203
 CC FT NP_BIND 268 271
 CC FT MOD_RES 177 177
 CC FT MOD_RES 350 350
 CC ADP-RIBOSYL(1) (BY ACTION OF CTX).
 CC ADP-RIBOSYL(1) (BY ACTION OF IAP).
 CC SQ SEQUENCE 353 AA; 40407 MW; 8BE2763493DE0F9 CRC64;
 CC
 CC Query Match 96.2%; Score 25; DB 1; Length 353;
 CC Best Local Similarity 83.3%; Pred. No. 39;
 CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 TIGSVI 6
 CC ID GBK RAT STANDARD; PRT; 353 AA.
 CC AC P08753;
 CC DT 01-NOV-1988 (Rel. 09, Created)
 CC DT 01-OCT-1994 (Rel. 30, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Guanine nucleotide-binding protein G(K), alpha subunit (G(i) alpha-3).
 CC GN GNA13 OR GNA13.
 CC OS Rattus norvegicus (Rat).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;
 CC
 CC RN (1)
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=8819830; PubMed=2834384;
 CC RA Itoh H., Toyama R., Kozasa T., Tsukamoto T., Matsuo K., Kaziro Y.;
 RT "Presence of three distinct molecular species of G protein alpha
 CC subunit. Structure of rat cDNAs and human genomic DNA.";
 CC J. Biol. Chem. 265:6656-6664(1988).
 CC
 CC RN (2)
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=88007678; PubMed=2820999;
 CC RA Jones D.T., Reed R.R.;
 RT "Molecular cloning of five GTP-binding protein cDNA species from rat
 CC olfactory neuroepithelium.";
 CC J. Biol. Chem. 262:14241-14249(1987).
 CC
 CC -1- FUNCTION: Guanine nucleotide-binding proteins (G proteins) are
 CC involved as modulators or transducers in various transmembrane


```

CC CC ADENYLATE CYCLASE: THEY INHIBIT THE CYCLASE IN RESPONSE TO BETA-ADRENERGIC STIMULI.
CC CC -1- SUBUNIT: G proteins are composed of 3 units (alpha, beta and gamma).
CC CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC CC -1- TISSUE SPECIFICITY: MAINLY EXPRESSED IN THE BRAIN, LUNG AND KIDNEY.
CC CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1 (G1/O/T/Z)).
CC CC -----
CC CC THE SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcement/ or send an email to license@ebi.ac.uk).
CC CC -----
CC CC EMBL; D12132; BAA04764.1; -.
CC CC DR HSSP; P10824; IAS3.
CC CC DR InterPro; IPR001019; Gprotein_alpha.
CC CC DR Pfam; PF00503; G-alpha.1.
CC CC DR PRINTS; PR00318; GPROTEIN.
CC CC DR ProDom; PD000281; Gprotein_alpha.1.
CC CC DR SMART; SM00275; G_alpha.1.
CC CC KW GTP-binding; Transducer; ADP-ribosylation; Multigene family; Myristate; Palmitate; Lipoprotein.
CC CC FT INIT MET 0 BY SIMILARITY.
CC CC FT LEU ID 1 1 MYRISTATE (BY SIMILARITY).
CC CC FT LEU ID 2 2 PALMITATE (BY SIMILARITY).
CC CC FT NP_BIND 39 46 GTP (BY SIMILARITY).
CC CC FT NP_BIND 199 203 GTP (BY SIMILARITY).
CC CC FT NP_BIND 268 271 GTP (BY SIMILARITY).
CC CC FT MOD_RRS 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX).
CC CC FT MOD_RRS 350 350 ADP-RIBOSYL[1] (BY ACTION OF PTX).
CC CC SQ SEQUENCE 353 AA; 40250 MW; 04E8C5DFB82D979 CRC64;

Query Match 96.2%; Score 25; DB 1; Length 353;
Best Local Similarity 83.3%; Pred. No. 39;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIGSVI 6
   |||||
Db 76 TIGSVI 81

RESULT 8
GB1_CHICK STANDARD; PRT; 353 AA.
ID GB1_CHICK
AC P50146;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Guanine nucleotide-binding protein G(1), alpha-1 subunit (adenylate cyclase-inhibiting G alpha protein).
GN GNAL1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Archosauria; Aves; Neognathae; Galliformes; Phaethonidae; Phaethonidae; OC Gallus.
OX NCBI_TaxID=9031;
RX MEDLINE=95121926; PubMed=7821803;
RA Kilbourne E.J., Galper J.B.;
RT "Cloning of cDNAs coding for the G alpha 11 and G alpha 12 G-proteins from chick brain."
RL Gene 150:341-344(1994).
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEIN (G PROTEIN) ARE INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSDUCER SIGNALING SYSTEMS.
CC -1- FUNCTION: THE G(1) PROTEINS ARE INVOLVED IN HORMONAL REGULATION OF ADENYLATE CYCLASE: THEY INHIBIT THE CYCLASE IN RESPONSE TO BETA-ADRENERGIC STIMULI.

```

```

CC ADRENERGIC STIMULI.
CC -1- SUBUNIT: G proteins are composed of 3 units (alpha, beta and
CC gamma).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1
CC (G1/O/T/2/).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L24548; AAA5066.1; -.
CC PIR: I50237; I50237.
CC HSSP: P10824; IAS3.
CC InterPro: IPR001019; Gprotein_alpha.
CC Pfam: PF00503; G-alpha; 1.
CC PRINTS: PR00318; GPROTEINA.
CC SMART: SM00275; G_alpha; 1.
CC GTP-binding; Transducer; ADP-ribosylation; Multigene family;
CC Myristate; Palmitate; Lipoprotein.
CC KW INIT MET 0 BY SIMILARITY.
CC FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
CC FT LIPID 2 2 PALMITATE (BY SIMILARITY).
CC FT NP_BIND 39 46 GTP (BY SIMILARITY).
CC FT NP_BIND 199 203 GTP (BY SIMILARITY).
CC FT NP_BIND 268 271 GTP (BY SIMILARITY).
CC FT MOD_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX).
CC FT MOD_RES 350 350 ADP-RIBOSYL[1] (BY ACTION OF IAP).
CC SQ SEQUENCE 353 AA; 40247 MW; E1DD0C848140137C CRC64;

Query Match 96.2%; Score 25; DB 1; Length 353;
Best Local Similarity 83.3%; Pred. No. 39;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIOSVI 6
Db 76 TIOSII 81

RESULT 9
GB11 HUMAN STANDARD; PRT; 353 AA.
ID GB11 HUMAN P04898; P11015; P31871;
AC P04898; P11015; P31871;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Guanine nucleotide-binding protein G(i), alpha-1 subunit (Adenylylate
DE cyclase-inhibiting G alpha protein).
GN GNAIL.
OS Homo sapiens (Human), and
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_Taxid=9606, 9913;
RN [1]
RP SEQUENCE OF 1-100 FROM N.A.
RC SPECIES=Human;
RX MEDLINE=88198230; PubMed=2834384;
RA Itoh H., Toyama R., Kozasa T., Tsukamoto T., Matsuo M.,
RA Kaziro Y.;
RT "Presence of three distinct molecular species of Gi protein alpha
RT subunit. Structure of rat cDNAs and human genomic DNAs.";
RL J. Biol. Chem. 263:6656-6664 (1988).
RN [2]
RP SEQUENCE OF 5-353 FROM N.A.
RC SPECIES=Human;
RX MEDLINE=87260939; PubMed=3110783;
RA Bray P., Carter A., Guo V., Puckett C., Kamholz J., Spiegel A.,
RA Nirenberg M.;

```

```

RT "Human cDNA clones for an alpha subunit of Gi signal-transduction
RT protein.";
RT Proc. Natl. Acad. Sci. U.S.A. 84:5115-5119 (1987).
RN [3]
RN SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Brain;
RA Yu W., Gibbs R.A.;
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=Bovine;
RX MEDLINE=86136587; PubMed=2419165;
RA Mukunda T., Tanabe T., Takahashi H., Noda M., Haga K., Haga T.,
RA Ichihara A., Kangawa K., Hirayama M., Matsuo H., Numa S.;
RT "Primary structure of the alpha-subunit of bovine adenylylate cyclase-
RT inhibiting G-protein deduced from the cDNA sequence.";
RL FEBS Lett. 197:305-310 (1986).
RN [5]
RP SEQUENCE OF 105-353 FROM N.A.
RC SPECIES=Bovine;
RX MEDLINE=87017009; PubMed=3094012;
RA Michel T., Winslow J.W., Smith J.A., Seidman J.G., Neer E.J.;
RT "Molecular cloning and characterization of cDNA encoding the
RT GTP-binding protein alpha 1 and identification of a related protein,
RT alpha h.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:7663-7667 (1986).
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -1- FUNCTION: THE G(I) PROTEINS ARE INVOLVED IN HORMONAL REGULATION OF
CC ADENYLYLATE CYCLASE. THEY INHIBIT THE CYCLASE IN RESPONSE TO BETA-
CC ADRENERGIC STIMULI.
CC -1- SUBUNIT: G proteins are composed of 3 units (alpha, beta and
CC gamma).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1
CC (G1/O/T/2/).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M20596; AAA35893.1; -.
CC EMBL: M20594; AAA35893.1; JOINED.
CC EMBL: M20595; AAA35893.1; JOINED.
CC EMBL: M17219; AAA52581.1; -.
CC EMBL: AF055013; AAC09361.1; -.
CC EMBL: X03642; CAA27288.1; -.
CC EMBL: M14207; AAA30561.1; -.
CC PIR: A28318; RGHU11.
CC PIR: A28318; RGHU11.
CC PDB: 1KJY; 08-MAY-02.
CC Genew: HGNC:4384; GNAIL.
CC MIM: 138310; -.
CC GO: GO:0005866; C:plasma membrane; TAS.
CC GO: GO:0000263; F:heterotrimeric G-protein GTPase, alpha-subunit; TAS.
CC InterPro: IPR001019; Gprotein_alpha.
CC Pfam: PF00503; G-alpha; 1.
CC PRINTS: PR00318; GPROTEINA.
CC ProDom: PD000281; Gprotein_alpha; 1.
CC SMART: SM00275; G_alpha; 1.
CC GTP-binding; Transducer; ADP-ribosylation; Multigene family;
CC Myristate; Palmitate; Lipoprotein; 3D-structure.
CC KW INIT MET 0 BY SIMILARITY.
CC FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
CC FT LIPID 2 2 PALMITATE (BY SIMILARITY).
CC FT NP_BIND 39 46 GTP (BY SIMILARITY).
CC FT NP_BIND 199 203 GTP (BY SIMILARITY).
CC FT NP_BIND 268 271 GTP (BY SIMILARITY).

```

FT MOD_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX).
 FT MOD_RES 350 350 ADP-RIBOSYL[1] (BY ACTION OF IAP).
 FT CONFLICT 112 112 A -> S (IN REF. 5).
 FT CONFLICT 329 329 K -> N (IN REF. 5).
 FT CONFLICT 336 336 D -> E (IN REF. 5).
 FT SEQUENCE 353 AA; 40230 MW; B456C4E189530A6D CRC64;

Query Match 96.2%; Score 25; DB 1; Length 353;
 Best Local Similarity 83.3%; Pred. No. 39;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TIQSVI 6
 |||||
 Db 76 TIQSVI 81

RESULT 10

GB11_ORYLA STANDARD; PRT; 353 AA.
 AC P87383;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Guanine nucleotide-binding protein G(i), alpha-1 subunit (Adenylylate cyclase-inhibiting G alpha protein) (G(i) alpha subunit) (G(i) alpha a).
 GN GNAT1.
 OS Oryzias latipes (Medaka fish) (Japanese ricefish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 OC Belontiiformes; Adrianchthyidae; Oryziinae; Oryzias.
 NC NCB1_Taxid=8090;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Ovary;
 RX MEDLINE=98055713; PubMed=9395335;
 RA Oba Y., Yoshikuni M., Tanaka M., Mita M., Nagahama Y.;
 RT "Inhibitory guanine-nucleotide-binding-regulatory protein alpha subunit in medaka (Oryzias latipes) oocytes -- cDNA cloning and decreased expression of proteins during oocyte maturation.";
 RL Eur. J. Biochem. 249:846-853(1997).

-1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE SIGNALING SYSTEMS. THE G(I) PROTEINS ARE INVOLVED IN HORMONAL REGULATION OF ADENYLYLATE CYCLASE. THEY INHIBIT THE CYCLASE IN RESPONSE TO BETA-ADRENERGIC STIMULI.

-1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS, ALPHA, BETA AND GAMMA. THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.

-1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1 (G(I)/O/T/2).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AB001741; BAA19454.1; -.
 DR HSPF; P10824; IAS3.
 DR InterPro; IPR001019; Gprotein_alpha.
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PR00318; GPROTEIN.
 DR ProDom; PD000281; Gprotein_alpha; 1.
 DR SMART; SM00275; G_alpha; 1.
 KM GTP-binding; Transducer; ADP-ribosylation; Multigene family; Myristate; Palmitate; Lipoprotein.

KW INIT MET 0 BY SIMILARITY.
 FT LIPID 1 MYRISTATE (BY SIMILARITY).
 FT LIPID 2 PALMITATE (BY SIMILARITY).
 FT NP_BIND 39 GTP (BY SIMILARITY).

FT NP_BIND 199 203 GTP (BY SIMILARITY).
 FT NP_BIND 268 271 GTP (BY SIMILARITY).
 FT MOD_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX) (BY
 FT MOD_RES 350 350 ADP-RIBOSYL[1] (BY ACTION OF IAP) (BY
 FT SEQUENCE 353 AA; 40149 MW; FBD5A91D0D69DFC0 CRC64;

Query Match 96.2%; Score 25; DB 1; Length 353;
 Best Local Similarity 83.3%; Pred. No. 39;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TIQSVI 6
 |||||
 Db 76 TIQSVI 81

RESULT 11

GB11_RAT STANDARD; PRT; 353 AA.
 AC P10824;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Guanine nucleotide-binding protein G(i), alpha-1 subunit (Adenylylate cyclase-inhibiting G alpha protein).
 GN GNAT1 OR GNAT1-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=8807678; PubMed=2820999;
 RA Jones D.T., Reed R.R.;
 RT "Molecular cloning of five GTP-binding protein cDNA species from rat olfactory neuroepithelium.";
 RL J. Biol. Chem. 262:14241-14249(1987).

-1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE SIGNALING SYSTEMS. THE G(I) PROTEINS ARE INVOLVED IN HORMONAL REGULATION OF ADENYLYLATE CYCLASE. THEY INHIBIT THE CYCLASE IN RESPONSE TO BETA-ADRENERGIC STIMULI.

-1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS, ALPHA, BETA AND GAMMA. THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.

-1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1 (G(I)/O/T/2).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AB001741; BAA19454.1; -.
 DR HSPF; P10824; IAS3.
 DR InterPro; IPR001019; Gprotein_alpha.
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PR00318; GPROTEIN.
 DR ProDom; PD000281; Gprotein_alpha; 1.
 DR SMART; SM00275; G_alpha; 1.
 KM GTP-binding; Transducer; ADP-ribosylation; Multigene family; Myristate; Palmitate; Lipoprotein.

KW INIT MET 0 BY SIMILARITY.
 FT LIPID 1 MYRISTATE (BY SIMILARITY).
 FT LIPID 2 PALMITATE (BY SIMILARITY).
 FT NP_BIND 39 GTP (BY SIMILARITY).

```

RT bound state."
RL J. Biol. Chem. 273:21752-21758(1998).
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -1- FUNCTION: THE G(I) PROTEINS ARE INVOLVED IN HORMONAL REGULATION OF
CC ADENYLATE CYCLASE; THEY INHIBIT THE CYCLASE IN RESPONSE TO BETA-
CC ADRENERGIC STIMULI.
CC -1- SUBUNIT: G proteins are composed of 3 units (alpha, beta and
CC gamma).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1
CC (G(I/O/T/Z)).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M17527; AAA40825.1; -.
CC PIR: C27423; RGR11.
CC PDB: 1GIA; 30-SEP-94.
CC PDB: 1GIL; 14-FEB-95.
CC PDB: 1GIT; 12-FEB-97.
CC PDB: 1GGR; 15-MAY-95.
CC PDB: 1AGR; 16-JUN-97.
CC PDB: 1IGD; 27-NOV-95.
CC PDB: 1IG2; 12-FEB-97.
CC PDB: 1AS2; 12-NOV-97.
CC PDB: 1AS3; 12-NOV-97.
CC PDB: 1BH2; 25-NOV-98.
CC PDB: 1BOF; 16-FEB-99.
CC PDB: 1CIP; 09-APR-99.
CC PDB: 1IPR; 01-APR-99.
CC PDB: 1P05; 03-G-ALPHA.1.
CC PRINTS; PR000318; GPROTEIN_ALPHA.
CC DR SMART; SM00275; G_alpha; 1.
CC KM GTP-binding; Transducer; Multigene family; ADP-ribosylation;
CC MYristate; Palmitate; Lipoprotein; 3D-structure.
CC INIT_MER 0
CC LIPID 1
CC NP_BIND 39
CC NP_BIND 199
CC NP_BIND 268
CC MOD_RES 177
CC MOD_RES 350
CC STRAND 32
CC TURN 41
CC TURN 45
CC TURN 62
CC TURN 68
CC TURN 69
CC TURN 91
CC TURN 97
CC TURN 98
CC TURN 99
CC TURN 114
CC TURN 120
CC TURN 133
CC TURN 140
CC TURN 141
CC TURN 142
CC TURN 149
CC TURN 151
CC TURN 155
CC TURN 156
CC TURN 162
CC TURN 164
CC TURN 170
CC TURN 174

```

```

FT TURN 175 175
FT STRAND 183 190
FT TURN 191 192
FT STRAND 193 200
FT HELIX 204 213
FT TURN 215 216
FT STRAND 219 225
FT HELIX 226 230
FT STRAND 232 233
FT TURN 234 235
FT STRAND 236 240
FT HELIX 241 253
FT TURN 254 254
FT HELIX 256 258
FT TURN 259 260
FT STRAND 262 268
FT HELIX 270 277
FT TURN 278 279
FT HELIX 282 284
FT TURN 285 285
FT TURN 287 288
FT HELIX 285 307
FT TURN 308 309
FT TURN 313 315
FT STRAND 318 322
FT TURN 325 326
FT HELIX 328 345
SQ SEQUENCE 353 AA; 40214 MW; B23724E187E90A6D CRC64;

Query Match          96.2%; Score 25; DB 1; Length 353;
Best Local Similarity 83.3%; Pred. No. 39;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
DB 76 TIQSVI 81

RESULT 12
GB1L XENIA STANDARD; PRT; 353 AA.
ID GB1L XENIA
AC P27044;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Guanine nucleotide-binding protein G(i1), alpha-1 subunit (Adenylate
DE cyclase-inhibiting G alpha protein).
DE GN
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxId=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RX MEDLINE=90346157; PubMed=2116977;
RA Olate J., Martinez S., Purcell P., Jorquera H., Codina J.,
RA Birbaumer L., Allende J.E.;
RT "Molecular cloning and sequence determination of four different cDNA
RT species coding for alpha-subunits of G proteins from Xenopus laevis
RT oocytes."
RL FEBS Lett. 268:27-31(1990).
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -1- FUNCTION: THE G(I) PROTEINS ARE INVOLVED IN HORMONAL REGULATION OF
CC ADENYLATE CYCLASE; THEY INHIBIT THE CYCLASE IN RESPONSE TO BETA-
CC ADRENERGIC STIMULI.
CC -1- SUBUNIT: G proteins are composed of 3 units (alpha, beta and
CC gamma).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1

```

```

CC (G(I/O/T/Z)).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56089; CAA39569.1; -.
CC PIR; S11045; RGL11.
CC HSSP; P10824; IAS3.
CC InterPro; IPR001019; Gpotein_alpha.
CC Pfam; PF00503; G-alpha_1.
CC PRINTS; PR00318; GPROTEIN.
CC ProDom; PD000281; Gpotein_alpha; 1.
CC SMART; SM00275; G_alpha; 1.
CC GTP-binding; Transducer; ADP-riboylacton; Multigene family;
CC Myristate; Palmitate; Lipoprotein.
CC INIT_MET 0
CC By SIMILARITY.
CC FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
CC FT LIPID 2 2 PALMITATE (BY SIMILARITY).
CC FT NP_BIND 39 46 GTP (BY SIMILARITY).
CC FT NP_BIND 199 203 GTP (BY SIMILARITY).
CC FT NP_BIND 268 271 GTP (BY SIMILARITY).
CC FT MOD_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX).
CC FT MOD_RES 350 350 ADP-RIBOSYL[1] (BY ACTION OF IAP).
CC SQ SEQUENCE 353 AA; 40270 MW; 6B4E94F841B077D CRC64;

Query Match 96.2%; Score 25; DB 1; Length 353;
Best Local Similarity 83.3%; Pred. No. 39;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TIQSVI 6
Db 76 TIQSVI 81

RESULT 13
PRTP_HSV6Z STANDARD; PRT; 410 AA.
AC P525F4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable processing and transport protein (Fragment).
GN U40 OR RAL1.
OS Human herpesvirus (type 6 / strain Z29) (HHV6).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Roseolovirus.
OX NCBI_TaxID=36351;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95074921; PubMed=7983761;
RA Stamey F.R.; Dominguez G.; Black J.B.; Dambaugh T.R.; Pellet P.E.;
RT "intragenomic linear amplification of human herpesvirus 6B orlyt
RT suggests acquisition of orlyt by transposition.";
RL J. Virol. 69:589-596(1995).
CC -1- FUNCTION: THIS PROTEIN MAY AFFECT TRANSLLOCATION OF THE VIRUS
CC GLYCOPROTEINS TO MEMBRANS. IT IS INVOLVED IN CAPSID MATURATION
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES PRTP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L16947; AAB06338.1; -.

```

```

DR InterPro; IPR000501; Proc_transpt.
DR Pfam; PF01366; PRTP; 1.
KW Capsid assembly.
FT NON_TER 410
SQ SEQUENCE 410 AA; 46812 MW; 708CE54DE3B8FBD3 CRC64;

Query Match 96.2%; Score 25; DB 1; Length 410;
Best Local Similarity 83.3%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TIQSVI 6
Db 247 TIQSVI 252

RESULT 14
VP5_AHSV6 STANDARD; PRT; 504 AA.
ID VP5_AHSV6
AC 071026;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Outer capsid protein VP5.
GN S6 OR M6.
OS African horse sickness virus 6 (AHSV-6) (African horse sickness virus
OS (serotype 6)).
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX NCBI_TaxID=86060;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98278331; PubMed=9617769;
RA Williams C.F.; Inoue T.; Lucas A.-M.; Zanotto P.; Roy P.;
RT "The complete sequence of four major structural proteins of African
RT horse sickness virus serotype 6: evolutionary relationships within
RT and between the orbiviruses.";
RL Virus Res. 53:53-73(1998).
CC -1- FUNCTION: THE VP5 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VP2)
CC WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID.
CC -1- SIMILARITY: BELONGS TO THE REOVIRUSES VP5 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF021237; AAC40996.1; -.
DR InterPro; IPR000145; Orbl_VP5.
DR Pfam; PF00901; Orbl_VP5; 1.
KW Coat protein.
SQ SEQUENCE 504 AA; 56901 MW; 6962AB8060A1A672 CRC64;

Query Match 96.2%; Score 25; DB 1; Length 504;
Best Local Similarity 83.3%; Pred. No. 56;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TIQSVI 6
Db 55 TIQSVI 60

RESULT 15
VP5_AHSV9 STANDARD; PRT; 505 AA.
ID VP5_AHSV9
AC 096597;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Outer capsid protein VP5.
GN S6 OR M6.
OS African horse sickness virus 9 (AHSV-9) (African horse sickness virus

```

```

OS (serotype 9)).
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX NCBI_TaxID=10897;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=97189517; PubMed=9037735;
RA du Plessis M., Nel L.H.;
RT "Comparative sequence analysis and expression of the M6 gene, encoding
RT the outer capsid protein VP5, of African horsesickness virus serotype
RT nine."
RL Virus Res. 47:41-49(1997).
CC -!- FUNCTION: THE VP5 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VP2)
CC WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID.
CC -!- SIMILARITY: BELONGS TO THE REOVIRUSES VP5 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; U74489; AAB17570.1; -.
CC DR InterPro; IPR000145; Orbl_VP5.
CC DR Pfam; PF00901; Orbl_VP5; I.
CC KW Coat protein.
CC SQ SEQUENCE 505 AA; 56771 MW; 4E55590F4D45DD3D CRC64;

```

```

Query Match          96.2%; Score 25; DB 1; Length 505;
Best Local Similarity 83.3%; Pred. No. 56;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TIQSVI 6
   |||:|
Db 55 TIQSII 60

```

Search completed: November 26, 2003, 12:31:13
 Job time : 3.89157 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:25:40 ; Search time 5.56627 Seconds
(without alignments)
103.662 Million cell updates/sec

Title: US-09-230-111c-9
Perfect score: 26
Sequence: 1 TIQSVI 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*

1: PIR1:
2: PIR2:
3: PIR3:
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	153	2 B64640	acyl carrier prote
2	26	100.0	15281	2 S41309	cyclosporin synthe
3	25	96.2	140	2 T28918	hypothetical prote
4	25	96.2	213	2 A71049	hypothetical prote
5	25	96.2	304	2 A64904	probable fibrial
6	25	96.2	304	2 C90892	probable adhesin (
7	25	96.2	304	2 F85725	aspartate semialde
8	25	96.2	329	2 B89916	hypothetical prote
9	25	96.2	342	2 T34450	hypothetical prote
10	25	96.2	345	1 RGLI3	hypothetical prote
11	25	96.2	354	1 RGHU11	hypothetical prote
12	25	96.2	354	1 RGHU11	hypothetical prote
13	25	96.2	354	1 RGHU11	hypothetical prote
14	25	96.2	354	1 RGHU11	hypothetical prote
15	25	96.2	354	1 RGHU11	hypothetical prote
16	25	96.2	354	1 RGHU11	hypothetical prote
17	25	96.2	354	1 RGHU11	hypothetical prote
18	25	96.2	354	1 RGHU11	hypothetical prote
19	25	96.2	354	1 RGHU11	hypothetical prote
20	25	96.2	354	1 RGHU11	hypothetical prote
21	25	96.2	354	1 RGHU11	hypothetical prote
22	25	96.2	354	1 RGHU11	hypothetical prote
23	25	96.2	354	1 RGHU11	hypothetical prote
24	25	96.2	354	1 RGHU11	hypothetical prote
25	25	96.2	354	1 RGHU11	hypothetical prote
26	25	96.2	354	1 RGHU11	hypothetical prote
27	25	96.2	354	1 RGHU11	hypothetical prote
28	25	96.2	354	1 RGHU11	hypothetical prote
29	25	96.2	354	1 RGHU11	hypothetical prote

30	24	92.3	194	2 T28889	hypothetical prote
31	24	92.3	209	2 T44507	meiz2 protein (imp
32	24	92.3	248	2 S05007	calcium channel al
33	24	92.3	296	2 C95961	probable sugar upc
34	24	92.3	301	2 S51439	hypothetical prote
35	24	92.3	360	2 I53032	bone morphogenetic
36	24	92.3	398	2 AE1879	hypothetical prote
37	24	92.3	458	2 A95893	hypothetical prote
38	24	92.3	472	1 BMH3	probable aminotran
39	24	92.3	476	2 JC4646	bone morphogenetic
40	24	92.3	478	2 JC4838	bone morphogenetic
41	24	92.3	487	2 T18952	hypothetical prote
42	24	92.3	500	2 T39385	probable mitochon
43	24	92.3	572	2 T34273	hypothetical prote
44	24	92.3	632	2 T21602	hypothetical prote
45	24	92.3	659	2 E64594	penicillin-binding

ALIGNMENTS

RESULT 1
B64640
acyl carrier protein - Helicobacter pylori (strain 26695)
C.Species: Helicobacter pylori
C.Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 26-May-2000
C.Accession: B64640
R.Tomb, J.F., White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D. Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khajek, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, L. Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: B64640
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-153 <TOM>
A:Cross-references: GB:AE000605; GB:AE000511; NID:g2314103; PIDN:AAD08009.1; PID:g2314103
C:Superfamily: acyl carrier protein homology
C:Keywords: carrier protein
F:78-148/Domain: acyl carrier protein homology <ACP>

Query Match 100.0%; Score 26; DB 2; Length 153;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
DB 83 TIQSVI 88

RESULT 2
S41309
cyclosporin synthetase - cyclosporin fungus
C.Species: Tolypocladium inflatum (cyclosporin fungus)
C.Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 03-Nov-2000
C.Accession: S45487; S41309
R.Weber, G.; Schoerndorfer, K.; Schneider-Scherzer, E.; Lechner, E. Curr. Genet. 26, 120-125, 1994
A:Title: The peptide synthetase catalyzing cyclosporine production in Tolypocladium nive
A:Reference number: S45487; MUID:95094306; PMID:8001164
A:Status: nucleic acid sequence not shown
A:Accession: S45487
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-15281 <WEB>
A:Cross-references: EMBL:Z28383; NID:g440168; PIDN:CAA62227.1; PID:g440169
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
A>Note: only a part of the translation is shown
A>Note: the source is designated as Tolypocladium inflatum
C:Genetics:
A:Genetic code: SGC3
C:Superfamily: cyclosporin fungus cyclosporin synthetase; acetate-CoA ligase homology; a

C.Keywords: carrier protein; phosphopantetheine; phosphoprotein
F.532-1008/Domain: acetate-CoA ligase homology <ACLI>
F.1029-1096/Domain: acyl carrier protein homology <ACP1>
F.1618-2069/Domain: acetate-CoA ligase homology <ACLI2>
F.2527-2594/Domain: acyl carrier protein homology <ACP2>
F.3115-3553/Domain: acetate-CoA ligase homology <ACLI3>
F.4014-4081/Domain: acyl carrier protein homology <ACP3>
F.4601-5057/Domain: acetate-CoA ligase homology <ACLI4>
F.5506-5573/Domain: acyl carrier protein homology <ACP4>
F.6099-6546/Domain: acetate-CoA ligase homology <ACLI5>
F.7003-7070/Domain: acyl carrier protein homology <ACP5>
F.7591-8042/Domain: acetate-CoA ligase homology <ACLI6>
F.8063-8130/Domain: acyl carrier protein homology <ACP6>
F.8659-9625/Domain: acetate-CoA ligase homology <ACLI7>
F.9558-9625/Domain: acyl carrier protein homology <ACP7>
F.10146-10586/Domain: acetate-CoA ligase homology <ACLI8>
F.11055-11122/Domain: acyl carrier protein homology <ACP8>
F.11635-12106/Domain: acetate-CoA ligase homology <ACLI9>
F.12127-12194/Domain: acyl carrier protein homology <ACP9>
F.12715-13159/Domain: acetate-CoA ligase homology <ACLI10>
F.13623-13690/Domain: acyl carrier protein homology <ACP10>
F.14213-14676/Domain: acetate-CoA ligase homology <ACLI11>
F.14698-14765/Domain: acyl carrier protein homology <ACP11>
F.1060,2558,4045,5537,7034,8094,9589,11086,12158,13654,14729/Binding site: phosphopantetheine
Query Match 100.0%; Score 26; DB 2; Length 15281;
Best Local Similarity 100.0%; Pred.No. 2.9e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TIQSVI 6
DB 3017 TIQSVI 3022
RESULT 3
28918
hypothetical protein C13F10.2 - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jun-2000
C.Accession: T28918
R.Tin, A.; Wohlmann, P.
submitted to the EMBL Data Library, April 1997
A.Description: The sequence of C. elegans cosmid C13F10.
A.Reference number: Z20543
A.Accession: T28918
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-140 <TIN>
A.Cross-references: EMBL:U97006; PIDN:AAC47964.1; GSPDB:GN00023; CESP:C13F10.2
A.Experimental source: strain Bristol N2; clone C13F10
C.Genetics:
A.Gene: CESP:C13F10.2
A.Map position: 5
A.Introns: 31/3; 59/2; 127/1
C.Superfamily: Caenorhabditis elegans hypothetical protein C13F10.2
Query Match 96.2%; Score 25; DB 2; Length 140;
Best Local Similarity 83.3%; Pred.No. 32;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TIQSVI 6
DB 48 TIQSVI 53
RESULT 4
A71049
hypothetical protein PH1681 - Pyrococcus horikoshii
C.Species: Pyrococcus horikoshii
C.Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C.Accession: A71049
R.Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998
A.Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A.Reference number: A71000; MUID:98344137; PMID:9679194
A.Accession: A71049
A.Status: preliminary; nucleic acid sequence not shown; translation not shown
A.Molecule type: DNA
A.Residues: 1-213 <KAW>
A.Cross-references: GB:AP000006; NID:g3236133; PIDN:BAJ0763.1; PID:g3258110
A.Experimental source: strain OT3
A.Note: this accession replaces an interim accession for a sequence replaced by GenBank
C.Genetics:
A.Gene: PH1681
Query Match 96.2%; Score 25; DB 2; Length 213;
Best Local Similarity 83.3%; Pred.No. 51;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TIQSVI 6
DB 43 TIQSVI 48
RESULT 5
A64904
probable fimbrial protein b1502 - Escherichia coli (strain K-12)
C.Species: Escherichia coli
C.Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C.Accession: A64904
R.Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A.Title: The complete genome sequence of Escherichia coli K-12.
A.Reference number: A64720; MUID:97426617; PMID:9276503
A.Accession: A64904
A.Status: nucleic acid sequence not shown; translation not shown
A.Molecule type: DNA
A.Residues: 1-304 <BLAT>
A.Cross-references: GB:AE000247; GB:U00096; NID:g1787773; PIDN:AAC74575.1; PID:g1787779,
A.Experimental source: strain K-12, substrain MG1655
C.Superfamily: fimbrial protein fimb
C.Keywords: fimbria
Query Match 96.2%; Score 25; DB 2; Length 304;
Best Local Similarity 83.3%; Pred.No. 75;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TIQSVI 6
DB 292 TIQSVI 297
RESULT 6
C90892
probable adhesin [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C.Species: Escherichia coli
C.Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C.Accession: C90892
R.Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A.Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A.Reference number: A99629; MUID:21156231; PMID:11258796
A.Accession: C90892
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-304 <HAY>
A.Cross-references: GB:BA000007; PIDN:BA835530.1; PID:g13361573; GSPDB:GN00154
A.Experimental source: strain O157:H7, substrain RIMD 0509952
C.Genetics:
A.Gene: EC82107
C.Superfamily: fimbrial protein fimb
Query Match 96.2%; Score 25; DB 2; Length 304;

Best Local Similarity 83.3%; Pred. No. 75;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
|:|:|:|
DB 292 TIQSVI 297

RESULT 7

F85725

probable adhesin, fimb type protein Z2206 [imported] - Escherichia coli (strain O157:H7,
C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: F85725

R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.D.; Mayhew
Iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimantanta, E.; Potamouelis, K.; Apodaca,
Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A05480; MUID:21074935; PMID:11206551

A:Accession: F85725

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-304 <STO>

A:Cross-references: GB:A005174; NID:g12515169; PIDN:AGS6266.1; GSPDB:GN00145; UMGF:222

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetic8:

A:Gene: Z2206

C:Superfamily: fimbrial protein fimb

Query Match 96.2%; Score 25; DB 2; Length 304;
Best Local Similarity 83.3%; Pred. No. 75;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
|:|:|:|
DB 292 TIQSVI 297

RESULT 8

B89316

aspartate semialdehyde dehydrogenase [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: B89316

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogun
ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A09750; MUID:21311952; PMID:11418146

A:Accession: B89316

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-329 <KUR>

A:Cross-references: GB:BA000018; PID:g13701191; PIDN:BA042486.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetic8:

A:Gene: asd

C:Superfamily: aspartate-semialdehyde dehydrogenase

Query Match 96.2%; Score 25; DB 2; Length 329;
Best Local Similarity 83.3%; Pred. No. 82;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
|:|:|:|
DB 128 TIQSVI 133

RESULT 9

T34450

hypothetical protein T19H12.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T34450

R:Davidson, S.

submitted to the EMBL Data Library, April 1997

A:Description: The sequence of C. elegans cosmid T19H12.

A:Reference number: 221528

A:Accession: T34450

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-342 <DAV>

A:Cross-references: EMBL:U97009; PIDN:AAC69029.1; GSPDB:GN00023; CESP:T19H12.4

A:Experimental source: strain Bristol N2; clone T19H12

C:Genetics:

A:Gene: CESP:T19H12.4

A:Map position: 5

A:Introns: 17/1; 84/2; 116/1; 254/3

Query Match 96.2%; Score 25; DB 2; Length 342;
Best Local Similarity 83.3%; Pred. No. 85;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 10

RGXLI3

GTP-binding regulatory protein G1 alpha-3 chain (adenylate cyclase-inhibiting) - African

N:Alternate names: GTP-binding regulatory protein Gk alpha chain; guanine nucleotide bin

C:Species: Xenopus laevis (African clawed frog)

C>Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001

C:Accession: S11046

R:Olarte, J.; Martinez, S.; Parcell, P.; Jorquera, H.; Codina, J.; Birnbaumer, L.; Allen

FEBS Lett. 268, 27-31, 1990

A:Title: Molecular cloning and sequence determination of four different cDNA species co

A:Reference number: S11045; MUID:90346157; PMID:2116977

A:Accession: S11046

A:Molecule type: mRNA

A:Residues: 1-345 <OLA>

A:Cross-references: GB:X56090; NID:g64709; PIDN:CA39570.1; PID:g64710

A:Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay

signals. The beta and gamma chains, required for GTPase activity, appear to be common to al

phae. It is specific for each type of G protein.

C:Comment: The G1 alpha chain is specific for G protein that is involved in hormonal re

C:Superfamily: GTP-binding regulatory protein Gs alpha chain

C:Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transduction

F:31-38/Region: nucleotide-binding motif A (P-loop)

F:260-263/Region: GTP-binding NKXD motif

F:337/Binding site: GTP (GTP) #status predicted

F:169/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

F:342/Modified site: ADP-ribosylcysteine (Cys) (by pertussis toxin) #status predicted

Query Match 96.2%; Score 25; DB 1; Length 345;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
|:|:|:|
DB 68 TIQSVI 73

RESULT 11
GTP-binding regulatory protein G1 alpha-1 chain (adenylate cyclase-inhibiting) - human

N:Alternate names: guanine nucleotide binding protein G1 alpha-1 chain; heterotrimeric C

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1992 #sequence_revision 22-Nov-1996 #text_change 19-Jan-2001

C:Accession: A28318; D28154; T08669

R:Bray, P.; Carter, A.; Guo, V.; Puckett, C.; Kamholz, J.; Spiegel, A.; Nirenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 84, 5115-5119, 1987

A:Title: Human cDNA clones for an alpha subunit of G1 signal-transduction protein.

A:Reference number: A28318; MUID:87260339; PMID:3110783

A/Accession: A28318
 A/Molecule type: mRNA
 A/Residues: 6-354 <BRA>
 A/Cross-references: GB:MI7219; NID:q183410; PIDN:AAA2581.1; PID:g386747
 R:Itou, H.; Toyama, R.; Kozasa, T.; Tanamoto, T.; Matsuka, M.; Kaziro, Y.
 A/Biol. Chem. 263, 6656-6664, 1988
 A/Title: Presence of three distinct molecular species of G-1 protein alpha-subunit. *Seru*
 A/Reference number: A28154; MUID:88198220; PMID:2834384
 A/Accession: D28154
 A/Status: not compared with conceptual translation
 A/Molecule type: DNA
 A/Residues: 1-101 <ITO>
 A/Cross-references: GB:M20596; GB:M19476; NID:q183189; PIDN:AAA35893.1; PID:g183191
 R:Duetschhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 A/Reference number: Z16467
 A/Reference number: Z16467
 A/Accession: T08669
 A/Molecule type: mRNA
 A/Residues: 'MGC5AATGSAATVPRDGPQTRDGLASRAGKQSLVVRNSRPLLSAPLRTASPTPLRKWGRGRPREAHE
 A/Cross-references: EMBL:AL049933
 A/Experimental source: fetal brain; clone DKFZP564K1216
 A/Note: differences are due to different assignment of start codons
 C/Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay
 ams. The beta and gamma chains, required for GTPase activity, appear to be common to al
 l; it is specific for each type of G protein.
 C/Comment: The G1 alpha chain is specific for G protein that is involved in hormonal reg
 C/Genetics:
 A/Genes: GDB:GNM11
 A/Cross-references: GDB:120001; OMIM:139310
 A/Map position: 7q21-7q21
 A/Note: DKFZP564K1216.1
 C/Superfamily: GTP-binding regulatory protein Gs alpha chain
 C/Keywords: blocked amino end; GTP binding; heterotrimer; lipoprotein; myristylation; nu
 F/2-354/Product: GTP-binding regulatory protein G1 alpha-1 chain #status predicted <MAY>
 F/40-47/Region: nucleotide-binding motif A (P-loop)
 F/269-272/Region: GTP-binding NKXD motif
 F/2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
 F/3/Binding site: palmitate (Cys) (covalent) #status predicted
 F/178/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted
 F/351/Modified site: ADP-ribosylcysteine (Cys) (by pertussis toxin) #status predicted

Query Match 96.2%; Score 25; DB 1; Length 354;
 Best Local Similarity 83.3%; Pred. No. 88;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
 |||||
 Db 77 TIQSVI 82

RESULT 12

RGR011

GTP-binding regulatory protein G1 alpha-1 chain (adenylate cyclase-inhibiting) - bovine
 N/Alternate names: guanine nucleotide binding protein G1 alpha-1 chain; heterotrimeric G
 C/Species: Bos primigenius taurus (cattle)
 C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #ext_change 19-Jan-2001
 C/Accession: A23631; A25888
 R:Nukada, T.; Tanabe, T.; Takahashi, H.; Noda, M.; Haga, K.; Haga, T.; Ichiyama, A.; Kan
 FEBS Lett. 197, 305-310, 1986
 A/Title: Primary structure of the alpha-subunit of bovine adenylate cyclase-inhibiting G
 A/Reference number: A23631; MUID:86136587; PMID:2419165
 A/Accession: A23631
 A/Molecule type: mRNA
 A/Residues: 1-354 <NUK>
 A/Cross-references: GB:X03642; NID:q390; PIDN:CAA27288.1; PID:g391
 R:Michel, T.; Winslow, J.W.; Smith, J.A.; Seidman, J.G.; Neer, E.J.
 Proc. Natl. Acad. Sci. U.S.A. 83, 7663-7667, 1986
 A/Title: Molecular cloning and characterization of cDNA encoding the GTP-binding protein
 A/Reference number: A94131; MUID:87017009; PMID:3094012
 A/Accession: A25888
 A/Molecule type: mRNA
 A/Residues: 106-112; 'S', 114-329; 'N', 331-336; 'E', 338-354 <MIC>
 A/Cross-references: GB:M14207; NID:g163129; PIDN:AAA30561.1; PID:g163130

C/Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay
 ams. The beta and gamma chains, required for GTPase activity, appear to be common to al
 l; it is specific for each type of G protein.
 C/Comment: The G1 alpha chain is specific for G protein that is involved in hormonal reg
 C/Superfamily: GTP-binding regulatory protein Gs alpha chain
 C/Keywords: blocked amino end; GTP binding; heterotrimer; lipoprotein; myristylation; nu
 F/2-354/Product: GTP-binding regulatory protein G1 alpha-1 chain #status predicted <MAY>
 F/40-47/Region: nucleotide-binding motif A (P-loop)
 F/269-272/Region: GTP-binding NKXD motif
 F/2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
 F/3/Binding site: palmitate (Cys) (covalent) #status predicted
 F/178/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted
 F/351/Modified site: ADP-ribosylcysteine (Cys) (by pertussis toxin) #status predicted

Query Match 96.2%; Score 25; DB 1; Length 354;
 Best Local Similarity 83.3%; Pred. No. 88;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
 |||||
 Db 77 TIQSVI 82

RESULT 13

RGR011

GTP-binding regulatory protein G1 alpha-1 chain (adenylate cyclase-inhibiting) - rat
 N/Alternate names: guanine nucleotide binding protein G1 alpha-1 chain; heterotrimeric C
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #ext_change 19-Jan-2001
 C/Accession: C27423
 R:Jones, D.T.; Reed, R.R.
 J. Biol. Chem. 262, 14241-14249, 1987
 A/Title: Molecular cloning of five GTP-binding protein cDNA species from rat olfactory r
 A/Reference number: A92614; MUID:88007678; PMID:2820599
 A/Accession: C27423
 A/Molecule type: mRNA
 A/Residues: 1-354 <JON>
 A/Cross-references: GB:M17527; NID:q203167; PIDN:AAA40825.1; PID:g203168
 C/Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay
 ams. The beta and gamma chains, required for GTPase activity, appear to be common to al
 l; it is specific for each type of G protein.
 C/Comment: The G1 alpha chain is specific for G protein that is involved in hormonal reg
 C/Superfamily: GTP-binding regulatory protein Gs alpha chain
 C/Keywords: blocked amino end; GTP binding; heterotrimer; lipoprotein; myristylation; nu
 F/2-354/Product: GTP-binding regulatory protein G1 alpha-1 chain #status predicted <MAY>
 F/40-47/Region: nucleotide-binding motif A (P-loop)
 F/269-272/Region: GTP-binding NKXD motif
 F/2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
 F/3/Binding site: palmitate (Cys) (covalent) #status predicted
 F/178/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted
 F/351/Modified site: ADP-ribosylcysteine (Cys) (by pertussis toxin) #status predicted

Query Match 96.2%; Score 25; DB 1; Length 354;
 Best Local Similarity 83.3%; Pred. No. 88;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
 |||||
 Db 77 TIQSVI 82

RESULT 14

RGR011

GTP-binding regulatory protein G1 alpha-1 chain (adenylate cyclase-inhibiting) - African
 N/Alternate names: guanine nucleotide binding protein G1 alpha-1 chain; heterotrimeric C
 C/Species: Xenopus laevis (African clawed frog)
 C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #ext_change 19-Jan-2001
 C/Accession: S11045
 R:Olave, J.; Martinez, S.; Purcell, P.; Jorquera, H.; Codina, J.; Birnbaumer, L.; Allen
 FEBS Lett. 268, 27-31, 1990
 A/Title: Molecular cloning and sequence determination of four different cDNA species co
 A/Reference number: S11045; MUID:90346157; PMID:2116977
 A/Accession: S11045

A.Molecule type: mRNA
 A.Residues: 1-354 <OLA>
 A.Cross-references: GB:J56089; NID:964707; PIDN:CAA35969.1; PID:964708
 C.Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay
 ains. The beta and gamma chains, required for GTPase activity, appear to be common to all
 rase; it is specific for each type of G protein.
 C.Comment: The G1 alpha chain is specific for G protein that is involved in hormonal reg
 C.Superfamily: GTP-binding regulatory protein Gs alpha chain
 C.Keyword: blocked amino end; GTP binding; heterotrimer; 11poprotein; myristylation; m
 F:2-354/Product: GTP-binding regulatory protein G1 alpha-1 chain #status predicted <MATS
 F:40-47/Region: nucleotide-binding motif A (P-loop)
 F:269-272/Region: GTP-binding NKXD motif
 F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
 F:3/Binding site: palmitate (Cys) (covalent) #status predicted
 F:178/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted
 F:351/Modified site: ADP-ribosylcysteine (Cys) (by pertussis toxin) #status predicted

Query Match 96.2%; Score 25; DB 1; Length 354;
 Best Local Similarity 83.3%; Pred. No. 88;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIGSVI 6
 |||||
 Db 77 TIGSII 82

RESULT 15
 RGHU13
 GTP-binding regulatory protein G1 alpha-3 chain (adenylate cyclase-inhibiting) - human
 N.Alternate names: GTP-binding regulatory protein Gk alpha chain; guanine nucleotide bin
 C.Species: Homo sapiens (man)
 C.Date: 31-Mar-1992 #sequence revision 31-Mar-1992 #ext change 19-Jan-2001
 C.Accession: S02348; S00055; A32139; C28154; A28157; S00078
 R.Beale, C.R.; Wilson, C.B.; Perlmutter, R.M.
 Proc. Natl. Acad. Sci. U.S.A. 84, 7886-7890, 1987
 A.Title: A small multigene family encodes G(i) signal-transduction proteins.
 A.Reference number: S02319; MUID:88068503; PMID:3120178
 A.Accession: S02348
 A.Molecule type: mRNA
 A.Residues: 1-354 <BEA>
 A.Cross-references: EMBL:J03005; NID:g183183; PIDN:AAA52557.1; PID:g183184
 R.Didbury, J.R.; Snyderman, R.
 FEBS Lett. 219, 259-263, 1987
 A.Title: Molecular cloning of a new human G protein. Evidence for two G(i-alpha)-like pr
 A.Reference number: S00055; MUID:87247315; PMID:3109953
 A.Accession: S00055
 A.Molecule type: mRNA
 A.Residues: 1-354 <DID>
 A.Cross-references: EMBL:M27543; NID:g183395; PIDN:AAA52579.1; PID:g183396
 A.Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 17-Met
 R.Kim, S.; Ang, S.L.; Bloch, D.B.; Bloch, K.D.; Kawahara, Y.; Tolman, C.; Lee, R.; Seidm
 Proc. Natl. Acad. Sci. U.S.A. 85, 4153-4157, 1988
 A.Title: Identification of cDNA encoding an additional alpha-subunit of a human GTP-bind
 A.Reference number: A32139; MUID:88247980; PMID:3132707
 A.Accession: A32139
 A.Molecule type: mRNA
 A.Residues: 1-354 <KIM>
 A.Cross-references: GB:J03238; NID:g183686; PIDN:AAA5939.1; PID:g106822
 R.Itoh, H.; Toyama, R.; Kozasa, T.; Tsukamoto, T.; Matsuo, M.; Kaziro, Y.
 J. Biol. Chem. 263, 6656-6664, 1988
 A.Title: Presence of three distinct molecular species of G-1 protein alpha-subunit. Stru
 A.Reference number: A28154; MUID:88198250; PMID:2834384
 A.Accession: C28154
 A.Molecule type: DNA
 A.Residues: 1-354 <ITO>
 A.Cross-references: GB:J03220
 R.Codina, J.; Olate, J.; Abramowitz, J.; Mattern, R.; Cook, R.G.; Birnbaumer, L.
 J. Biol. Chem. 263, 6746-6750, 1988
 A.Title: Alpha-1-3 cDNA encodes the alpha-subunit of G-iK, the stimulatory G protein of r
 A.Reference number: A28157; MUID:88198244; PMID:2452165
 A.Accession: A28157
 A.Molecule type: mRNA
 A.Residues: 1-354 <COD>

A.Cross-references: GB:J03198; NID:g183224; PIDN:AAA5996.1; PID:g183225
 R.Suki, W.N.; Abramowitz, J.; Mattern, R.; Codina, J.; Birnbaumer, L.
 FEBS Lett. 220, 187-192, 1987
 A.Title: The human genome encodes at least three non-allelic G proteins with alpha(i)-1
 A.Reference number: S00078; MUID:87276552; PMID:2440724
 A.Accession: S00078
 A>Status: nucleic acid sequence not shown
 A.Molecule type: mRNA
 A.Residues: 1-354 <SUK>
 C.Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay
 ains. The beta and gamma chains, required for GTPase activity, appear to be common to a
 rase; it is specific for each type of G protein.
 C.Comment: The G1 alpha chain is specific for G protein that is involved in hormonal reg
 C.Genetics:
 A.Gene: GDB:GNAT3
 A.Cross-references: GDB:119276; OMIM:139370
 A.Map position: 1p13-1p13
 C.Superfamily: GTP-binding regulatory protein Gs alpha chain
 C.Keyword: blocked amino end; GTP binding; heterotrimer; 11poprotein; myristylation; m
 F:2-354/Product: GTP-binding regulatory protein G1 alpha-3 chain #status predicted <MAT
 F:40-47/Region: nucleotide-binding motif A (P-loop)
 F:269-272/Region: GTP-binding NKXD motif
 F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
 F:3/Binding site: palmitate (Cys) (covalent) #status predicted
 F:178/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted
 F:351/Modified site: ADP-ribosylcysteine (Cys) (by pertussis toxin) #status predicted

Query Match 96.2%; Score 25; DB 1; Length 354;
 Best Local Similarity 83.3%; Pred. No. 88;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIGSVI 6
 |||||
 Db 77 TIGSII 82

Search completed: November 26, 2003, 12:35:57
 Job time : 7.56627 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:25:10 ; Search time 13.0843 Seconds
(without alignments)
118.334 Million cell updates/sec

Title: US-09-230-111c-9
Perfect score: 26
Sequence: 1 TIQSVI 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virine:*
16: sp_bacterioplasmid:*
17: sp_archaeal:

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	153	025615	025615 helicobacte
2	26	100.0	316	091109	091109 streptococ
3	26	100.0	15281	009164	009164 tolypoclad
4	25	96.2	85	090713	090713 anaerobact
5	25	96.2	91	066001	066001 oenococcus
6	25	96.2	140	001488	001488 caenorhabd
7	25	96.2	202	090714	090714 anaerobact
8	25	96.2	209	028945	028945 sus scrofa
9	25	96.2	213	059320	059320 pyrococcus
10	25	96.2	258	086007	086007 oceanobact
11	25	96.2	265	090706	090706 uncultured
12	25	96.2	304	080402	080402 escherichia
13	25	96.2	323	001609	001609 caenorhabd
14	25	96.2	329	090713	090713 staphylococ
15	25	96.2	329	090713	090713 staphylococ
16	25	96.2	329	080402	080402 staphylococ

17	25	96.2	329	16	080402	080402 staphylococ
18	25	96.2	354	4	090404	090404 homo sapien
19	25	96.2	354	4	080402	080402 homo sapien
20	25	96.2	354	13	090846	090846 gallus gall
21	25	96.2	354	13	090847	090847 gallus gall
22	25	96.2	362	2	056591	056591 vibrio angu
23	25	96.2	380	16	080387	080387 raietonia s
24	25	96.2	527	12	080387	080387 raietonia s
25	25	96.2	527	12	080387	080387 raietonia s
26	25	96.2	527	12	080387	080387 raietonia s
27	25	96.2	527	12	080387	080387 raietonia s
28	25	96.2	527	12	080387	080387 raietonia s
29	25	96.2	559	16	090407	090407 pasteurella
30	25	96.2	600	5	080401	080401 encephalito
31	25	96.2	606	8	090407	090407 rhinophus
32	25	96.2	606	8	090407	090407 rhinophus
33	25	96.2	726	12	091809	091809 human herpe
34	25	96.2	726	12	091809	091809 human herpe
35	25	96.2	1010	5	080310	080310 manduca sex
36	25	96.2	1010	5	080310	080310 manduca sex
37	25	96.2	1236	5	070346	070346 plasmodium
38	25	96.2	1236	5	070346	070346 plasmodium
39	25	96.2	1285	4	090893	090893 caenorhabd
40	25	96.2	1285	4	090893	090893 caenorhabd
41	25	96.2	145	8	080407	080407 polytomella
42	25	96.2	194	5	022478	022478 caenorhabd
43	25	96.2	201	5	090407	090407 caenorhabd
44	25	96.2	208	2	090407	090407 escherichia
45	25	96.2	209	2	090407	090407 bacillus sp

ALIGNMENTS

RESULT 1
ID 025615 PRELIMINARY: PRT: 153 AA.
AC 025615;
DT 01-JAN-1998 (TRMBLrel. 05, Created)
DT 01-JAN-1998 (TRMBLrel. 05, Last sequence update)
DT 01-OCT-2002 (TRMBLrel. 22, Last annotation update)
DE Acyl carrier protein (ACP).
GN HP0962.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteriaceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=9794467; PubMed=9252185;
RA Tomb J.-F., White O., Kierlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
pylori."
RT Nature 388:539-547 (1997).
CC 1 - FUNCTION: THIS PROTEIN IS THE CARRIER OF THE GROWING FATTY ACID
CHAIN IN FATTY ACID BIOSYNTHESIS (BY SIMILARITY).
DR EMBL: A000605; AAD08009.1; -
DR HSSP: P02901; IACP.
DR TIGR: HP0962; -
DR InterPro: IPR003231; Acyl_carrier.
DR InterPro: IPR006163; Pp_bind.
DR Pfam: PF00550; pp-binding; 1.
DR ProDom: PD000887; Acyl_carrier; 1.
DR TIGRfam: TIGR00517; acyl_carrier; 1.

DR PROSITE; PS50075; ACP DOMAIN; 1.
KW Fatty acid biosynthesis; Hypothetical protein; Phosphopantetheine;
SQ COMPLETE PROTEOME.
SEQUENCE 153 AA; 17917 MW; 5A1D715D9A394404 CRC64;

Query Match 100.0%; Score 26; DB 16; Length 153;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
DB 83 TIQSVI 88

RESULT 2

Q9L1U9 PRELIMINARY; PRT; 316 AA.
AC Q9L1U9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Putative sugar transporter inner membrane protein.
GN SC02945 OR SCE59.04C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
NCBI_Taxid=1902;

SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Oliver K., Harris D.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cereno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

[3]

SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA MEDLINE=97000351; PubMed=843436;
RA Redenbach M., Kleiser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered contigs and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).

[4]

SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=2196410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleiser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleiser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinovitch E., Rajandream M.A., Rubinfeld K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2)." ;
RL Nature 417:141-147(2002).
DR EMBL; AL939114; CAB72191.1; -;
DR InterPro; IPR000515; BPD_transp.
DR Pfam; PF00528; BPD_transp; 1.
KW Sugar transport; Complete proteome.
SQ SEQUENCE 316 AA; 34367 MW; F4A637F5665DDFF CRC64;

Query Match 100.0%; Score 26; DB 16; Length 316;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6

DB 240 TIQSVI 245

RESULT 3
Q09164 PRELIMINARY; PRT; 15281 AA.
ID Q09164;
AC Q09164;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Cyclosporin synthetase (CYSYN) (EC 6.-.-.-).
GN SIMA.

OS Tolyposciadium infatum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Clavicipitaceae; mitosporic Clavicipitaceae;
OC Tolyposciadium.
NCBI_Taxid=29910;

SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=ATCC 34921;
RX MEDLINE=95094306; PubMed=8001164;
RA Weber G., Schoergerdoffer K., Schneider-Scherzer E., Leitner E.;
RT "The peptide synthetase catalyzing cyclosporine production in
Tolyposciadium niveum is encoded by a giant 45.8-kilobase open reading
frame." ;
RT Curr. Genet. 26:120-125(1994).

-I- FUNCTION: THE CONSTITUENT AMINO ACIDS OF CYCLOSPORINS ARE
ACTIVATED AS AMINOACYL-ADENYLATES WITH PEPTIDE BONDS FORMED
THROUGH THE PARTICIPATION OF AMINO ACID THIOLESTER INTERMEDIATES.

-I- COFACTOR: CONTAINS 11 COVALENTLY BOUND PHOSPHOPANTHEINES.

-I- PATHWAY: NON-RIBOSOMAL BIOSYNTHESIS OF CYCLOSPORINS.

-I- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT
COVALENT BINDING OF AMP TO THEIR SUBSTRATE.

DR EMBL; Z28363; CAA82227.1; -;
DR HSRP; P14687; LAMU.
DR InterPro; IPR000873; AMP-bind.
DR InterPro; IPR001242; Condensatn.
DR InterPro; IPR006162; Ppantne_attach.
DR InterPro; IPR006163; Pp_bind.
DR InterPro; IPR000511; SAM_bind.
DR Pfam; PF00501; AMP-binding; 11.
DR Pfam; PF00668; Condensation; 13.
DR Pfam; PR00550; pp-binding; 11.
DR PROSITE; PS50075; ACP_DOMAIN; 11.
DR PROSITE; PS00455; AMP_BINDING; 10.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; 9.
KW Ligase; Antibiotic biosynthesis; Phosphopantetheine; Repeat;
KW Multifunctional enzyme.

FT DOMAIN 15179 15219 13 X 3 AA APPROXIMATE REPEATS.
FT REPEAT 3 1086 DOMAIN 1.
FT REPEAT 1087 2585 DOMAIN 2.
FT REPEAT 2586 4072 DOMAIN 3.
FT REPEAT 4073 5564 DOMAIN 4.
FT REPEAT 5565 7061 DOMAIN 5.
FT REPEAT 7062 8121 DOMAIN 6.
FT REPEAT 8122 9616 DOMAIN 7.
FT REPEAT 9617 11113 DOMAIN 8.
FT REPEAT 11114 12185 DOMAIN 9.
FT REPEAT 12186 13681 DOMAIN 10.
FT REPEAT 13682 14767 DOMAIN 11 (ALA-ACTIVATING).
FT BINDING 1060 1060 PHOSPHOPANTHEINE (POTENTIAL).
FT BINDING 2558 2558 PHOSPHOPANTHEINE (POTENTIAL).
FT BINDING 4045 4045 PHOSPHOPANTHEINE (POTENTIAL).
FT BINDING 5537 5537 PHOSPHOPANTHEINE (POTENTIAL).
FT BINDING 7034 7034 PHOSPHOPANTHEINE (POTENTIAL).
FT BINDING 8094 8094 PHOSPHOPANTHEINE (POTENTIAL).
FT BINDING 9589 9589 PHOSPHOPANTHEINE (POTENTIAL).
FT BINDING 11086 11086 PHOSPHOPANTHEINE (POTENTIAL).
FT BINDING 12158 12158 PHOSPHOPANTHEINE (POTENTIAL).
FT BINDING 13654 13654 PHOSPHOPANTHEINE (POTENTIAL).
FT BINDING 14729 14729 PHOSPHOPANTHEINE (POTENTIAL).

SQ SEQUENCE 15281 AA; 1689051 MW; E26DA7AA35324C05 CRC64;

Query Match 100.0%; Score 26; DB 3; Length 15281;

Best Local Similarity 100.0%; Pred. No. 6e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
|:||||
3017 TIQSVI 3022

RESULT 4

ID 09U713 PRELIMINARY; PRT; 85 AA.

AC 09U713; 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

GN Parasitism-specific protein PSP10.

OS Anastrepha suspensa (Caribbean fruit fly).

OC Neoptera; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Tephritidae; Tephritidae; Anastrepha.

OX NCBI_TaxID=28587;

RN [1] SEQUENCE FROM N.A.

RP MEDLINE=99440192; PubMed10510497;

RA Shi X., Gomez S.P., Lawrence P.O.;

RT "A 24 kDa parasitism-specific protein from the Caribbean fruit fly,

RT Anastrepha suspensa: cDNA and deduced amino acid sequence.";

RL Insect Biochem. Mol. Biol. 29:749-755(1999).

DR EMBL; AF165877; AAF01363.1;

SQ SEQUENCE 85 AA; 9761 MW; 696F31348213D4AD CRC64;

QY 1 TIQSVI 6
|:||||
25 TIQSVI 30

RESULT 5

ID 066001 PRELIMINARY; PRT; 91 AA.

AC 066001; 01-AUG-1998 (TREMBlrel. 07, Created)

DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

DE 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

GN HDcB (Fragment).

OS Oenococcus oeni (Leuconostoc oenos).

OC Bacteria; Firmicutes; Lactobacillales; Oenococcus.

OX NCBI_TaxID=1247;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=9204;

RA MEDLINE=98397285; PubMed=9633629;

RT "Histidine carboxylase of Leuconostoc oenos 9204: purification,

RT kinetic properties, cloning and nucleotide sequence of the hcd gene.";

RL J. Appl. Microbiol. 84:143-151(1998).

DR EMBL; U58665; AAC38299.1;

FT NON TER 91 91

SQ SEQUENCE 91 AA; 9909 MW; AF4E009CA3C6A10B CRC64;

QY 1 TIQSVI 6
|:||||
5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query Match 96.2%; Score 25; DB 2; Length 91;

Best Local Similarity 83.3%; Pred. No. 52;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 11 TIQSVI 16
|:||||

RESULT 6

ID 001488 PRELIMINARY; PRT; 140 AA.

AC 001488; 01-JUL-1997 (TREMBlrel. 04, Created)

DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)

DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

GN Hypothetical 16.3 kDa protein.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Pelodierinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=Bristol N2;

RA MEDLINE=99069613; PubMed=9851916;

RT "Genome sequence of the nematode C. elegans: a platform for

RT investigating biology. The C. elegans Sequencing Consortium.";

RL Science 282:2012-2018(1998).

RN [2] SEQUENCE FROM N.A.

RP STRAIN=Bristol N2;

RA "The sequence of C. elegans cosmid C13F10.";

RT Submitted (Apr-1997) to the EMBL/Genbank/DBJ databases.

RN [3] SEQUENCE FROM N.A.

RP STRAIN=Bristol N2;

RA "Direct Submission.";

RT Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL; U97006; AAC47964.1;

SQ SEQUENCE 140 AA; 16271 MW; ACDA68370DD56A1E CRC64;

QY 1 TIQSVI 6
|:||||
48 TIQSVI 53

RESULT 7

ID 09U714 PRELIMINARY; PRT; 202 AA.

AC 09U714; 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

GN Parasitism-specific protein PSP24.

OS Anastrepha suspensa (Caribbean fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OX NCBI_TaxID=28587;

RN [1] SEQUENCE FROM N.A.

RP MEDLINE=99440192; PubMed=10510497;

RA Shi X., Gomez S.P., Lawrence P.O.;

RT "A 24 kDa parasitism-specific protein from the Caribbean fruit fly,

RT Anastrepha suspensa: cDNA and deduced amino acid sequence.";

RL Insect Biochem. Mol. Biol. 29:749-755(1999).

DR EMBL; AF165876; AAF01362.1;

FT NON TER 202 202

SQ SEQUENCE 202 AA; 21009 MW; AF4E009CA3C6A10B CRC64;

QY 1 TIQSVI 6
|:||||
5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query Match 96.2%; Score 25; DB 5; Length 140;

Best Local Similarity 83.3%; Pred. No. 81;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query Match 96.2%; Score 25; DB 2; Length 265;
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
 |||||
 Db 192 TIQSVI 197

RESULT 12
 O8XAX2 PRELIMINARY; PRT; 304 AA.
 AC O8XAX2;
 DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE Putative adhesin, similar to FimH protein.
 GN 22206 OR EC62107.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
 RA MEDLINE=21074935; PubMed=1120651;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 Rose J.G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potomsis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / RIMD 050952;
 RA MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shingawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22 (2001).
 DR EMBL; AF005354; AAG56266.1; -;
 DR EMBL; AP02557; BAB35530.1; -;
 DR InterPro; IPR000259; FimHrel.
 DR Pfam; PF00419; FimHrel; 1.
 GN Complete proteome.
 KW SEQUENCE 304 AA; EED538023D95AFD5 CRC64;

Query Match 96.2%; Score 25; DB 16; Length 304;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
 |||||
 Db 292 TIQSVI 297

RESULT 13
 O01609 PRELIMINARY; PRT; 323 AA.
 AC O01609;
 DT 01-JUL-1997 (TEMBLrel. 04, Created)
 DT 01-OCT-2001 (TEMBLrel. 18, Last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
 DE Hypothetical 37.0 kDa protein.
 GN T19H12.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Brictol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Brictol N2;
 RA Davidson S.;
 RT "The sequence of C. elegans cosmid T19H12.";
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Brictol N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U97009; AAC69029.2; -;
 DR WormPep; T19H12.4; CR26002.
 DR InterPro; IPR000168; 7TM_nematode.
 DR InterPro; IPR000276; GPCF_Rhodopsin.
 DR PROSITE; PS0282; G_PROTEIN_RECEP_F1_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 323 AA; 36972 MW; 076426D964B705F2 CRC64;

Query Match 96.2%; Score 25; DB 5; Length 323;
 Best Local Similarity 83.3%; Pred. No. 2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
 |||||
 Db 238 TIQSVI 243

RESULT 14
 O9EZ13 PRELIMINARY; PRT; 329 AA.
 AC O9EZ13;
 DT 01-MAR-2001 (TEMBLrel. 16, Created)
 DT 01-OCT-2001 (TEMBLrel. 22, Last sequence update)
 DE Aspartate semialdehyde dehydrogenase.
 GN ASD.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8325-4;
 RX MEDLINE=21340416; PubMed=11447207;
 RA Wilentz M.D., Foster S.J.;
 RT "Identification and Analysis of Staphylococcus aureus Components
 RT Expressed by a Model System of Growth in Serum.";
 RL Infect. Immun. 69:5198-5202 (2001).
 DR EMBL; AF306669; AAG42245.1; -;
 DR InterPro; IPR005986; Asp_ADH_USG1.
 DR InterPro; IPR000534; Semialdh_ch.
 DR Pfam; PF01118; Semialdh_ch; 1.
 DR Pfam; PF02774; Semialdh_ch; 1.
 DR TIGRFAMs; TIGR01296; asd B; 1.
 SQ SEQUENCE 329 AA; 36283 MW; 417FA1BC2811C7F5 CRC64;

Query Match 96.2%; Score 25; DB 2; Length 329;
 Best Local Similarity 83.3%; Pred. No. 2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIQSVI 6
 |||||

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:26:20 ; Search time 14.4578 Seconds
(without alignments)
102.059 Million cell updates/sec

Title: US-09-230-111c-11

Perfect score: 37

Sequence: 1 RETRIEVED 8

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 673684 seqs, 184443283 residues

Total number of hits satisfying chosen parameters: 673684

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	100.0	8	US-08-681-219-13	Sequence 13, Appl
2	37	100.0	8	US-10-092-138-11	Sequence 11, Appl
3	37	100.0	8	US-08-230-111C-11	Sequence 11, Appl
4	30	81.1	193	US-10-032-201B-62	Sequence 62, Appl
5	30	81.1	289	US-09-738-626-4130	Sequence 4130, Ap
6	30	81.1	295	US-10-144-156-18	Sequence 18, Appl
7	29	78.4	116	US-09-876-348A-4	Sequence 4, Appl
8	29	78.4	116	US-09-876-348A-4	Sequence 4, Appl
9	29	78.4	134	US-09-876-348A-3	Sequence 3, Appl
10	29	78.4	134	US-09-876-348A-3	Sequence 3, Appl
11	29	78.4	149	US-09-876-348A-27	Sequence 27, Appl
12	29	78.4	149	US-09-876-348A-27	Sequence 27, Appl
13	29	78.4	174	US-09-876-348A-25	Sequence 25, Appl
14	29	78.4	174	US-09-876-348A-25	Sequence 25, Appl
15	29	78.4	199	US-09-764-870-489	Sequence 489, Appl

16	29	78.4	199	US-09-764-853-850	Sequence 850, App
17	29	78.4	199	US-09-839-825-28	Sequence 28, Appl
18	29	78.4	199	US-10-125-540-489	Sequence 489, App
19	29	78.4	199	US-10-103-513-578	Sequence 578, App
20	29	78.4	157	US-10-156-761-13213	Sequence 13213, A
21	29	78.4	1139	US-10-080-170-460	Sequence 460, App
22	28	75.7	67	US-10-007-770-13	Sequence 13, Appl
23	28	75.7	149	US-09-738-626-6793	Sequence 6793, Ap
24	28	75.7	316	US-10-042-894A-25	Sequence 25, Appl
25	28	75.7	426	US-09-815-242-11282	Sequence 11282, A
26	28	75.7	443	US-10-156-761-12374	Sequence 12374, A
27	28	75.7	682	US-10-238-075-478	Sequence 478, App
28	28	75.7	692	US-09-815-242-11568	Sequence 11568, A
29	28	75.7	709	US-10-156-761-12453	Sequence 12453, A
30	28	75.7	798	US-10-007-270-9	Sequence 9, Appl
31	28	75.7	1269	US-09-998-027-2	Sequence 2, Appl
32	28	75.7	1269	US-10-165-099-2	Sequence 2, Appl
33	28	75.7	1339	US-10-082-830-282	Sequence 282, App
34	28	75.7	2008	US-09-736-969A-2	Sequence 2, Appl
35	28	75.7	2008	US-09-736-969A-91	Sequence 91, Appl
36	28	75.7	2008	US-09-736-960-88	Sequence 88, Appl
37	28	75.7	2008	US-09-736-968A-105	Sequence 105, App
38	28	75.7	2008	US-09-978-244A-29	Sequence 29, Appl
39	28	75.7	2073	US-09-978-244A-10	Sequence 10, Appl
40	27	73.0	16	US-09-765-614B-27	Sequence 27, Appl
41	27	73.0	16	US-09-925-715-23	Sequence 23, Appl
42	27	73.0	24	US-10-029-386-32648	Sequence 32648, A
43	27	73.0	50	US-10-221-161-183	Sequence 183, App
44	27	73.0	63	US-09-864-761-34296	Sequence 34296, A
45	27	73.0	231	US-10-006-922-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-08-681-219-13
Sequence 13, Application US/08681219
Publication No. US20020058607A1
GENERAL INFORMATION:
APPLICANT: Takaaki Sato and Junn Yanagisawa
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLG
TITLE OF INVENTION: (PDZ/DHR) DOMAIN AND USES THEREOF
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/681,219
FILING DATE: 22-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/48962/JPM/JKM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-681-219-13

Query Match 100.0%; Score 37; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RETESTV 8
Db 1 RETESTV 8

RESULT 2

US-10-092-138-11
Sequence 11, Application US/10092138
Publication No. US20030170723A1
GENERAL INFORMATION:
APPLICANT: Sato, Taka-Aki
TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
FILE REFERENCE: 65823/JPW/PT
CURRENT APPLICATION NUMBER: US/10/092,138
CURRENT FILING DATE: 2002-09-06
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial
US-10-092-138-11

Query Match 100.0%; Score 37; DB 12; Length 8;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RETESTV 8
Db 1 RETESTV 8

RESULT 3

US-09-230-111C-11
Sequence 11, Application US/09230111C
Publication No. US20030203414A1
GENERAL INFORMATION:
APPLICANT: Sato, Taka-Aki
APPLICANT: Yanagisawa, Junn
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN
TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF (PDZ/DHR)
FILE REFERENCE: 48962-A-PCT-US
CURRENT APPLICATION NUMBER: US/09/230,111C
CURRENT FILING DATE: 1999-05-17
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial
US-09-230-111C-11

Query Match 100.0%; Score 37; DB 12; Length 8;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RETESTV 8

Db 1 RETESTV 8

RESULT 4
US-10-032-201B-62
Sequence 62, Application US/10032201B
Publication No. US20030167524A1
GENERAL INFORMATION:
APPLICANT: Van Rooijen, GJ's
APPLICANT: Deckers, Harm
APPLICANT: Helfetz, Peter Bernard
APPLICANT: Briggs, Steven
APPLICANT: Dalmia, Bipin Kumar
APPLICANT: Del Val, Greg
APPLICANT: Zapachinski, Steve
APPLICANT: Moloney, Maurice
TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF MULTIMERIC PROTEINS, AND RELATED
FILE REFERENCE: 38814 351B
CURRENT APPLICATION NUMBER: US/10/032,201B
CURRENT FILING DATE: 2001-12-19
NUMBER OF SEQ ID NOS: 313
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 62
LENGTH: 193
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-10-032-201B-62

Query Match 81.1%; Score 30; DB 12; Length 193;
Best Local Similarity 62.5%; Pred. No. 92;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RETESTV 8
Db 180 RETLEKTI 187

RESULT 5
US-09-738-626-4130
Sequence 4130, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 4130
LENGTH: 289
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-4130

Query Match 81.1%; Score 30; DB 10; Length 289;

Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RETESTV 8
|:|:|
Db 41 RETVELTV 48

RESULT 6
US-10-144-156-18

/ Sequence 18, Application US/10144156
/ Publication No. US20030166197A1
/ GENERAL INFORMATION:

/ APPLICANT: Ecker, Joseph R.
/ APPLICANT: Nehring, Ramiah
/ APPLICANT: McGrath, Robert B.
/ TITLE OF INVENTION: ETHYLENE INSENSITIVE PLANTS
/ FILE REFERENCE: SALKINS.040A
/ CURRENT APPLICATION NUMBER: US/10/144,156
/ CURRENT FILING DATE: 2002-05-10
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 18
/ LENGTH: 295
/ TYPE: PRT

/ ORGANISM: Caenorhabditis elegans
US-10-144-156-18

Query Match 81.1%; Score 30; DB 12; Length 295;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RETESTV 7
|:|:|
Db 60 RETELAT 66

RESULT 7
US-09-876-348A-4

/ Sequence 4, Application US/09876348A
/ Patent No. US20020172951A1
/ GENERAL INFORMATION:

/ APPLICANT: Horwath, K. L. and Myers, K. L.
/ TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
/ TITLE OF INVENTION: Antifreeze Proteins and Method for Assaying Activity.
/ FILE REFERENCE: RB-125-RI
/ CURRENT APPLICATION NUMBER: US/09/876,348A
/ CURRENT FILING DATE: 2001-08-09
/ PRIOR APPLICATION NUMBER: 60/210,446
/ PRIOR FILING DATE: 2000-06-08
/ NUMBER OF SEQ ID NOS: 48
/ SOFTWARE: Microsoft Word
/ SEQ ID NO 4
/ LENGTH: 116
/ TYPE: PRT

/ ORGANISM: Tenebrio molitor
/ OTHER INFORMATION: Mature Protein for Tm 13.17
US-09-876-348A-4

Query Match 78.4%; Score 29; DB 10; Length 116;
Best Local Similarity 62.5%; Pred. No. 84;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RETESTV 8
|:|:|
Db 91 RDTVEETV 98

RESULT 8
US-09-876-796A-4

/ Sequence 4, Application US/09876796A
/ Patent No. US20020173024A1
/ GENERAL INFORMATION:

/ APPLICANT: Horwath, K. L. and Easton, C. M.
/ TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
/ TITLE OF INVENTION: Antifreeze Proteins and Method for Assaying Activity.
/ FILE REFERENCE: RB-125-SEQ
/ CURRENT APPLICATION NUMBER: US/09/876,796A
/ CURRENT FILING DATE: 2001-06-07
/ PRIOR APPLICATION NUMBER: 60/210,446
/ PRIOR FILING DATE: 2000-06-08
/ NUMBER OF SEQ ID NOS: 48
/ SOFTWARE: Microsoft Word
/ SEQ ID NO 4
/ LENGTH: 116
/ TYPE: PRT

/ ORGANISM: Tenebrio molitor
/ OTHER INFORMATION: Mature Protein for Tm 13.17
US-09-876-796A-4

Query Match 78.4%; Score 29; DB 10; Length 116;
Best Local Similarity 62.5%; Pred. No. 84;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RETESTV 8
|:|:|
Db 91 RDTVEETV 98

RESULT 9
US-09-876-348A-3

/ Sequence 3, Application US/09876348A
/ Patent No. US20020172951A1
/ GENERAL INFORMATION:

/ APPLICANT: Horwath, K. L. and Myers, K. L.
/ TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
/ TITLE OF INVENTION: Antifreeze Proteins and Method for Assaying Activity.
/ FILE REFERENCE: RB-125-RI
/ CURRENT APPLICATION NUMBER: US/09/876,348A
/ CURRENT FILING DATE: 2001-08-09
/ PRIOR APPLICATION NUMBER: 60/210,446
/ PRIOR FILING DATE: 2000-06-08
/ NUMBER OF SEQ ID NOS: 48
/ SOFTWARE: Microsoft Word
/ SEQ ID NO 3
/ LENGTH: 134
/ TYPE: PRT

/ ORGANISM: Tenebrio molitor
/ OTHER INFORMATION: Precursor Protein for Tm 13.17
US-09-876-348A-3

Query Match 78.4%; Score 29; DB 10; Length 134;
Best Local Similarity 62.5%; Pred. No. 99;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RETESTV 8
|:|:|
Db 109 RDTVEETV 116

RESULT 10
US-09-876-796A-3

/ Sequence 3, Application US/09876796A
/ Patent No. US20020173024A1
/ GENERAL INFORMATION:

/ APPLICANT: Horwath, K. L. and Easton, C. M.
/ TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
/ TITLE OF INVENTION: Antifreeze Proteins and Method for Assaying Activity.
/ FILE REFERENCE: RB-125-SEQ
/ CURRENT APPLICATION NUMBER: US/09/876,796A
/ CURRENT FILING DATE: 2001-06-07
/ PRIOR APPLICATION NUMBER: 60/210,446
/ PRIOR FILING DATE: 2000-06-08
/ NUMBER OF SEQ ID NOS: 48
/ SOFTWARE: Microsoft Word
/ SEQ ID NO 3

```

; LENGTH: 134
; TYPE: PRT
; ORGANISM: Tenebrio molitor
; OTHER INFORMATION: Precursor Protein for Tm 13.17
US-09-876-796A-3

Query Match
Best Local Similarity 78.4%; Score 29; DB 10; Length 134;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETESTV 8
DB 109 RDTVEETV 116

RESULT 11
US-09-876-348A-27
; Sequence 27, Application US/09876348A
; Patent No. US20020172951A1
; GENERAL INFORMATION:
; APPLICANT: Horwath, K. L. and Myers, K. L.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
; FILE REFERENCE: RB-125-RI
; CURRENT APPLICATION NUMBER: US/09/876,348A
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/210,446
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Word
; SEQ ID NO 27
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Tenebrio molitor
; OTHER INFORMATION: Mature Protein with His-tag, Tm 13.17
US-09-876-348A-27

Query Match
Best Local Similarity 78.4%; Score 29; DB 10; Length 149;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETESTV 8
DB 124 RDTVEETV 131

RESULT 12
US-09-876-796A-27
; Sequence 27, Application US/09876796A
; Patent No. US20020173024A1
; GENERAL INFORMATION:
; APPLICANT: Horwath, K. L. and Easton, C. M.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
; FILE REFERENCE: RB-125-SEQ
; CURRENT APPLICATION NUMBER: US/09/876,796A
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/210,446
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Word
; SEQ ID NO 27
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Tenebrio molitor
; OTHER INFORMATION: Mature Protein with His-tag, Tm 13.17
US-09-876-796A-27

Query Match
Best Local Similarity 78.4%; Score 29; DB 10; Length 149;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETESTV 8
DB 124 RDTVEETV 131
```

```

DB 124 RDTVEETV 131

RESULT 13
US-09-876-348A-25
; Sequence 25, Application US/09876348A
; Patent No. US20020172951A1
; GENERAL INFORMATION:
; APPLICANT: Horwath, K. L. and Myers, K. L.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
; FILE REFERENCE: RB-125-RI
; CURRENT APPLICATION NUMBER: US/09/876,348A
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/210,446
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Word
; SEQ ID NO 25
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Tenebrio molitor
; OTHER INFORMATION: Precursor Protein with His-tag, Tm 13.17
US-09-876-348A-25

Query Match
Best Local Similarity 78.4%; Score 29; DB 10; Length 174;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETESTV 8
DB 149 RDTVEETV 156

RESULT 14
US-09-876-796A-25
; Sequence 25, Application US/09876796A
; Patent No. US20020173024A1
; GENERAL INFORMATION:
; APPLICANT: Horwath, K. L. and Easton, C. M.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
; FILE REFERENCE: RB-125-SEQ
; CURRENT APPLICATION NUMBER: US/09/876,796A
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/210,446
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Word
; SEQ ID NO 25
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Tenebrio molitor
; OTHER INFORMATION: Precursor Protein with His-tag, Tm 13.17
US-09-876-796A-25

Query Match
Best Local Similarity 78.4%; Score 29; DB 10; Length 174;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETESTV 8
DB 149 RDTVEETV 156

RESULT 15
US-09-764-870-489
; Sequence 489, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```

FILE REFERENCE: PT214
CURRENT APPLICATION NUMBER: US/09/764,870
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 646
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 489
LENGTH: 199
TYPE: prt
ORGANISM: Homo sapiens
US-09-764-870-489

Query Match 78.4%; Score 29; DB 9; Length 199;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Oy 1 RETEST 7
DB 157 RDTVEST 163

Search completed: November 26, 2003, 12:38:38
Job time : 14.4578 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 26, 2003, 12:21:29 ; Search time 22.1687 Seconds
(without alignments)
57.280 Million cell updates/sec

Title: US-09-230-111c-11

Perfect score: 37

Sequence: 1 RETESTV 8

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	100.0	8	19	AAW50178
2	31	83.8	22	AAW67502	Signal-transducing
3	30	81.1	193	AA37884	Propionibacterium
4	30	81.1	193	ABR60713	Arabidopsis thaliana
5	30	81.1	203	AA39661	Arabidopsis thaliana
6	30	81.1	217	AA37883	Arabidopsis thaliana
7	30	81.1	289	AA390376	C. glutamicum prote
8	30	81.1	295	ABP57742	C. elegans Y48B5A.
9	30	81.1	391	AAU63147	Propionibacterium

10	30	81.1	650	23	ABR93055
11	29	78.4	91	22	AAU62952
12	29	78.4	117	24	ABU71056
13	29	78.4	134	23	AAU10036
14	29	78.4	149	23	AAU10049
15	29	78.4	174	23	AAU76232
16	29	78.4	181	23	AAU10048
17	29	78.4	196	22	AAU64952
18	29	78.4	199	22	ABR10542
19	29	78.4	199	22	AAU18507
20	29	78.4	199	22	AAU43654
21	29	78.4	199	22	AAU07851
22	29	78.4	199	22	AAU19839
23	29	78.4	199	22	AAU21851
24	29	78.4	199	23	ABP67129
25	29	78.4	199	23	ABP48059
26	29	78.4	199	24	ABG74532
27	29	78.4	200	22	AAU18405
28	29	78.4	221	22	ABG06533
29	29	78.4	221	22	ABR32920
30	29	78.4	414	14	AAW65406
31	29	78.4	414	19	AAW65406
32	29	78.4	414	22	AAW67075
33	29	78.4	414	24	AAE34355
34	29	78.4	955	15	AAW57365
35	29	78.4	955	17	AAW03691
36	29	78.4	1139	23	ABU05809
37	29	78.4	3830	22	AAW97820
38	28	75.7	10	19	AAW57880
39	28	75.7	16	15	AAW44717
40	28	75.7	18	10	AAW90286
41	28	75.7	109	22	AAU31460
42	28	75.7	110	22	AAW78165
43	28	75.7	124	22	AAW78905
44	28	75.7	149	22	AAW93039
45	28	75.7	213	21	AAW70003

ALIGNMENTS

RESULT 1	AAW50178	standard; peptide; 8 AA.
ID	AAW50178	
AC	AAW50178	
XX		
DT	16-JUL-1998	(first entry)
XX		
DE	Signal-transducing protein carboxy-terminal peptide.	
XX		
KW	Inhibition; specific binding; signal-transducing protein;	
KM	cytoplasmic protein; proliferation; cancer cell; apoptosis;	
KW	virally infected cell.	
XX		
OS	Synthetic.	
XX		
PN	WO9805347-A1.	
XX		
PD	12-FEB-1998.	
XX		
PF	18-JUL-1997;	97WO-US12677.
XX		
PR	22-JUL-1996;	96US-0681219.
XX		
PA	(UYCO) UNIV COLUMBIA NEW YORK.	
XX		
PI	Sato T, Yanagisawa J;	
XX		
DR	WPI; 1998-145347/13.	
XX		
PT	Inhibition of signal transduction - by inhibiting binding between a	
PT	signal-transducing protein and a cytoplasmic protein, for treating	

Herbicidally activ
Propionibacterium
Human adipocyte Se
Tm13.17 anti-freez
Tm 13.17 clone/His
Tm 13.17 clone/His
Tm 13.17 clone/His
Shrimp white spot
Human cDNA SEQ ID
Human endocrine po
Human polypeptide
Novel human serine
Human novel extrac
Novel human neopla
Human polypeptide
Human polypeptide
Human Down syndrom
Human endocrine po
Novel human diagno
Sequence of a euk
Human transcriptio
Human shear stress
Human yin-yang 1 (K39 polypeptide of
Leishmania chagasi
M. tuberculosis an
Mouse spactin prot
Human milk beta-ca
DNA synthesis acce
DNA synthesis acce
Novel human bectet
Human tumour relat
C. glutamicum prote
Human beta-casein.

PT e.g. cancer or viral infection
XX
PS Claim 20; Page 64; 108bp; English.
CC
CC A novel composition is capable of inhibiting specific binding
CC between a signal-transducing protein (STP) having the
CC carboxy-terminal sequence (Ser/Thr)-Xaa-(Val/Ile/Leu), where Xaa =
CC any amino acid (e.g. the present peptide), and a cytoplasmic
CC protein (CP) containing the sequence AAW50162 or AAW50163.
CC The composition can be used to inhibit the proliferation of cancer
CC or virally infected cells, or induce apoptosis in cancer or virally
CC infected cells.
XX
SQ Sequence 8 AA;
Query Match 100.0%; Score 37; DB 19; Length 8;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RETESTV 8
| | | | |
Db 1 RETESTV 8
RESULT 2
AAU67502
ID AAU67502 standard; Protein; 81 AA.
XX
AC AAU67502;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #28398.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
XX 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR MPI: 2001-616774/71.
DR N-PSDB; AAS59529.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 28697; 1069bp; English.
XX
CC Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the

CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 81 AA;
Query Match 83.8%; Score 31; DB 22; Length 81;
Best Local Similarity 87.5%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 RETESTV 8
| | | | |
Db 24 RETESTV 31
RESULT 3
AAG37884
ID AAG37884 standard; Protein; 193 AA.
XX
AC AAG37884;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46656.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136682.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match
Best Local Similarity 81.1%; Score 30; DB 21; Length 193;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RETIESTV 8
Db 180 RETLEKTI 187

RESULT 4
ABP60713
ID ABP60713 standard; Protein; 193 AA.
XX
XX ABP60713;
XX
XX 06-SEP-2002 (first entry)

DE Arabidopsis thaliana thioedoxin M-type 4 SEQ ID NO:62.

KM Multimeric protein; redox protein; thioedoxin; thioedoxin reductase;
KM oil body; ophthalmological; antidiabetic; cytoskeletal; antiproliferative;
KM vasotrophic; valnerary; antibacterial; immunosuppressive; antitumor;
KM food product; milk; wheat; oxidative stress; cataract; diabetes;
KM chronic obstructive pulmonary disease; emphysema; psoriasis; sepsis;
KM bronchiolopulmonary disease; malignancy; reperfusion injury; wound healing;
KM gastro intestinal bleeding; intestinal bowel disease; ulcer;
KM gastro oesophageal reflux disease.

OS Arabidopsis thaliana.

XX MO200250289-A1.

XX 27-JUN-2002.

XX 19-DEC-2001; 2001WO-US0240.

XX 19-DEC-2000; 2000US-0742900.

XX 05-JUL-2001; 2001US-302885P.

XX 04-DEC-2001; 2001US-0006038.

XX (SEMB-) SEMBIOSYS GENETICS INC.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Van Rooijen G, Deckers H, Heifetz PB, Briggs SP, Dalmia BK;
PI Del Val G, Zaplachinski S, Moloney M;

XX WPI; 2002-508806/54.

PT Producing oil body associated with recombinant multimeric protein
PT complex e.g. redox proteins and immunoglobulins comprises producing
PT recombinant polypeptides capable of forming the complex in cells
PT comprising oil bodies -

XX Claim 81; Page 214-215; 362pp; English.

CC The present invention describes a method (M1) for producing an oil body
CC associated with a recombinant multimeric protein complex (MPC). M1
CC comprises producing in a cell comprising oil bodies a first and second
CC recombinant polypeptide (P1, P2), where P1 is capable of associating
CC with P2 to form the MPC and associating the complex with an occlusion
CC body (OB) through an OB-targeting-protein capable of associating with OB
CC and P1. M1 is useful for producing an oil body associated with a
CC recombinant MPC. The oil bodies are further formulated for use in the
CC preparation of a food product such as milk or wheat based food product,
CC personal care product which reduces the oxidative stress on the surface

CC area of the human body or used to lighten the skin, or a pharmaceutical
CC composition used to treat chronic obstructive pulmonary disease (COPD),
CC cataracts, diabetes, emphysema, bronchiolopulmonary disease, psoriasis,
CC malignancies, reperfusion injury, wound healing, sepsis, gastro
CC intestinal (GI) bleeding, intestinal bowel disease (IBD), ulcers, GERD
CC (gastro oesophageal reflux disease). ABN89569 to ABN89593 and ABP60677
CC to ABP60964 represent sequence given in the exemplification of the
CC present invention.

XX SQ Sequence 193 AA;

Query Match
Best Local Similarity 81.1%; Score 30; DB 23; Length 193;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RETIESTV 8
Db 180 RETLEKTI 187

RESULT 5
AAG39661
ID AAG39661 standard; Protein; 203 AA.
XX
XX AAG39661;
XX
XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 49105.

KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 30-APR-1999; 99US-0132407.

XX 04-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 06-MAY-1999; 99US-0132487.

XX 07-MAY-1999; 99US-0132863.

XX 11-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 14-MAY-1999; 99US-0134370.

XX 18-MAY-1999; 99US-0134768.

XX 19-MAY-1999; 99US-0134941.

XX 20-MAY-1999; 99US-0135124.

XX 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143544.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 81.1%; Score 30; DB 21; Length 203;
Best Local Similarity 62.5%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 RETESTV 8
DB 180 RETLEKTI 187
RESULT 6
AAG37883
ID AAG37883 standard; Protein; 217 AA.
XX
AC AAG37883;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46655.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PR 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145819.
PR 27-JUL-1999; 99US-0145819.
PR 28-JUL-1999; 99US-0145851.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.

```

PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 81.1%; Score 30; DB 21; Length 217;
 Best Local Similarity 62.5%; Pred. No. 2e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETESTV 8
 Db 204 RETLEKTI 211

RESULT 7
 AAG90376
 ID AAG90376 standard; Protein; 289 AA.

AC AAG90376;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 4130.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

KW organic acid synthesis.

OS Corynebacterium glutamicum.

XX EPI108790-A2.

PD 20-JUN-2001.

PF 18-DEC-2000; 2000EP-0127688.

XX 16-DEC-1999; 99UP-0377484.

PR 07-APR-2000; 2000UP-0159162.

PR 03-AUG-2000; 2000UP-0280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayaashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR WPI: 2001-376931/40.

DR N-PSDB; AAH65595.

PT Novel polynucleotides derived from Coryneform bacteria, for identifying

PT mutation point of a gene, measuring expression of a gene, analysing

PT expression profile or pattern of a gene and identifying homologous gene

PT Claim 17; SEQ ID NO: 4130; 246bp + Sequence Listing; English.

PS

XX The present invention provides a number of nucleotide and protein

CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These

CC are useful for identifying the mutation point of a gene derived from a

CC mutant of coryneform bacterium, measuring expression amount and

CC analysing the expression profile or expression pattern of a gene derived

CC from Coryneform bacterium, and identifying a homologue of a gene derived

CC from coryneform bacterium. Coryneform bacteria are useful for producing

CC amino acids, nucleic acids, vitamins, saccharides and organic acids,

CC particularly L-lysine. The present sequence is a protein described

CC in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from the

CC

XX Sequence 289 AA;

QY Query Match 81.1%; Score 30; DB 22; Length 289;
 Best Local Similarity 75.0%; Pred. No. 2.8e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETESTV 8
 Db 41 RETVELTV 48

RESULT 8
 AAB57742

```

ID   ABP57742 standard; Protein; 295 AA.
XX
AC   ABP57742;
XX
DT   24-JAN-2003 (first entry)
XX
DE   C. elegans Y48B6A.11.
XX
KW   EIN6; ethylene insensitive 6; ethylene; plant; delayed ripening;
RW   flowering; delayed senescence; browning; pathogen sensitivity; longevity;
KW   fruitfly; CG9088.
XX
OS   Caenorhabditis elegans.
XX
PN   WO200290540-A1.
PD   14-NOV-2002.
XX
PE   10-MAY-2002; 2002MO-US15049.
PR   10-MAY-2001; 2001US-290303P.
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
PI   Ecker JR, Nehring R, McGrath RB;
DR   WPI; 2003-067792/06.
XX
PT   New isolated EIN6 polynucleotide for conferring ethylene insensitive
PT   phenotype in plants, such as delayed ripening, flowering, delayed
PT   senescence or browning, which may improve quality and longevity of
PT   crops -
PS   Disclosure; Fig 7; 78pp; English.
XX
CC   The invention relates to a novel EIN6 (ethylene insensitive 6)
CC   polynucleotide. The polynucleotide is useful in conferring an ethylene
CC   insensitive phenotype in plants, such as delayed ripening, flowering,
CC   delayed senescence, browning, or altered sensitivity to pathogens, which
CC   may improve the quality and longevity of many crops. The present sequence
CC   represents Y48B6A.11, a predicted protein from C. elegans.
CX
SQ   Sequence      295 AA;
Query Match          81.1%; Score 30; DB 24; Length 295;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches    6; Conservative    1; Mismatches    0; Indels    0; Gaps    0;
OY         1 RETEST 7
           |||||:|
Db        60 RETREAT 66
RESULT 9
ID       AAU63147
AAU63147 standard; Protein; 391 AA.
XX
AC       AAU63147;
XX
DT       27-FEB-2002 (first entry)
XX
DE       Propionibacterium acnes immunogenic protein #24043.
XX
KW       SAPRO syndrome; synovitis; acne; pustulosis; hypetroiosis; osteomyelitis;
KW       uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW       inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW       dermatological; osteopathic; neuroprotectant.
XX
OS       Propionibacterium acnes.
XX
PN       WO200181581-A2.
PD       01-NOV-2001.

```

XX	20-APR-2001; 2001WO-US12865.
XX	
XX	
XX	21-APR-2000; 2000US-199047P.
PR	02-JUN-2000; 2000US-208841P.
PR	07-JUL-2000; 2000US-216747P.
XX	
PA	(CORI-) CORIYA CORP.
XX	
PI	Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhalaria A;
PI	L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX	
DR	WPI; 2001-616774/71.
DR	N-PsDB; AAS59632.
XX	
PT	Propionibacterium acnes polypeptides and nucleic acids useful for
PT	vaccinating against and diagnosing infections, especially useful for
PT	treating acne vulgaris -
XX	
PS	Example 1; SEQ ID No 24342; 1069pp; English.
XX	
XX	Sequences AAU93105-AAU68017 represent Propionibacterium acnes immunogenic
CC	polypeptides. The proteins and their associated DNA sequences are used in
CC	the treatment, prevention and diagnosis of medical conditions caused by
CC	P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC	pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC	P. acnes is also involved in infections of bone, joints and the central
CC	nervous system, however it is particularly involved in the inflammatory
CC	lesions associated with acne vulgaris. A method for detecting the
CC	presence or absence of P. acnes in a patient comprises contacting a
CC	sample with a binding agent that binds to the proteins of the invention.
CC	and determining the amount of bound protein in the sample. The
CC	polypeptides may be used as antigens in the production of antibodies
CC	specific for P. acnes proteins. These antibodies can be used to
CC	downregulate expression and activity of P. acnes polypeptides and
CC	therefore treat P. acnes infections. The antibodies may also be used as
CC	diagnostic agents for determining P. acnes presence, for example, by
CC	enzyme linked immunosorbent assay (ELISA).
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SEQ	Sequence 391 AA;
XX	
Query Match	81.1%; Score 30; DB 22; Length 391;
Best Local Similarity	62.5%; Pred. No. 4e+02;
Matches	5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY	1 RETESTV 8
DB	348 REAVESTI 355
XX	
RESULT 10	
ABR33055	ABR33055 standard; Protein; 650 AA.
XX	
AC	ABR33055;
XX	
DT	31-MAY-2002 (first entry)
XX	
DE	Herbicidally active polypeptide SEQ ID NO 2266.
XX	
KW	Herbicidal; plant; agriculture; herbicide.
XX	
OS	Arabidopsis thaliana.
XX	
PN	WO200210210-A2.
XX	
PD	07-FEB-2002.
XX	
PF	28-AUG-2001; 2001WO-EP09892.
XX	

XX The invention relates to a complex between two interacting proteins in
 CC adipocyte cells, given in the specification. The proteins are identified
 CC by selecting a bait protein from a known adipocyte marker and then
 CC performing a yeast 2-hybrid selection to isolate prey proteins encoded by
 CC members of an adipocyte cDNA library. The proteins are designated SID
 CC (RtM) (selected interacting domains) proteins. Also included are a
 CC polynucleotide encoding a polypeptide in the adipocyte cells, a
 CC recombinant host cell expressing at least one of the interacting
 CC polypeptides of the complex, selecting a modulating compound in adipocyte
 CC cells, a SID (RtM) polypeptide comprising any of the 738 amino acid
 CC sequences given in the specification (including its fragment or variant),
 CC a SID (RtM) polynucleotide comprising any of the 738 nucleotide sequences
 CC given in the specification (including its fragment or variant), a vector
 CC comprising the SID (RtM) polynucleotide, a recombinant host cell
 CC comprising the vector, a protein chip comprising the polypeptides and
 CC a record comprising all or part of the data, listed in the specification.
 CC The complex, polypeptides, polynucleotides and compounds are
 CC useful for preventing or treating metabolic disorders such as obesity
 CC or diabetes. The polynucleotides are useful as probes or primers. The
 CC complex is particularly useful for identifying selected interacting
 CC domains (SID (RtM)) for screening drugs that modulate the protein
 CC interaction, thus exhibiting the therapeutic effect. The present
 CC sequence represents a SID (prey) protein of the invention.

SQ Sequence 117 AA;

Query Match 78.4%; Score 29; DB 24; Length 117;
 Best Local Similarity 85.7%; Pred. No. 1.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ETESTV 8
 :|||||
 Db 54 QTESTV 60

RESULT 13
 AAU10036
 ID AAU10036 standard; Protein; 134 AA.
 AC AAU10036;
 XX
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Tm13.17 anti-freeze protein.
 XX
 XX Anti-freeze peptide; Tm 13.17; yellow mealworm beetle; THP; cryoinjury;
 KW thermal hysteresis protein; cryosurgery; cold climatisation;
 KW hypothermic cell preservation; de-icing formulation; transgenic plant;
 KW transgenic animal.
 OS Tenebrio molitor.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..18
 FT /note= "Signal peptide"
 FT Protein 19..116
 FT /note= "Mature Tm13.17 protein"
 XX
 XX WO200194378-A1.
 XX
 PD 13-DEC-2001.
 XX
 PF 07-JUN-2001; 2001WO-US18532.
 XX
 PR 08-JUN-2000; 2000US-210446P.
 XX
 XX (UNIV) UNIV NEW YORK STATE RES FOUND.
 PA (HORN/) HORNATH K L.
 PA (MYER/) MYERS K L.
 PA (EAST/) EASTON C M.
 XX
 PI Horwath KL, Myers KL, Easton CM;

XX
 DR WPI: 2002-090137/12.
 DR N-PsDB; AAs21109.
 XX
 PT New cDNA polynucleotide encoding a thermal hysteresis protein which is
 PT a Type III anti-freeze protein derived from the Tenebrionidea
 PT Superfamily, useful for providing antifreeze protection to improve the
 PT quality of food -
 XX
 PS Claim 15; Fig 2.6a; 364pp; English.

XX This invention relates to a cDNA polynucleotide comprising a nucleotide
 CC sequence encoding a thermal hysteresis protein (THP) which is a Type III
 CC anti-freeze protein derived from the Tenebrionidea Superfamily. Anti-
 CC freeze proteins lower the freezing point of a solution without affecting
 CC the melting point of the solution. An activated anti-freeze protein may
 CC be incorporated into plant, produce or fish in an amount sufficient to
 CC provide antifreeze protection or in a region of a target tissue to
 CC cryoinjury during cryosurgery. The proteins of the invention may also be
 CC used in hypothermic solutions or bathing media to reduce cold damage in
 CC order to provide cryogenic or hypothermic preservation of cells and
 CC tissues. The proteins may be used as de-icing formulations or used on
 CC surfaces to reduce existing ice buildup or abate the formation of ice
 CC buildup on surfaces such as a road, aircraft, household products,
 CC machinery and plant surfaces or as a food product to improve the quality
 CC of food by abating freezing of solutions, freezer burn, or degradation
 CC due to cold storage. The polynucleotides for the activated protein can
 CC be used to create transgenic or gene-modified plants, crops, fish, or
 CC animals having greater tolerance to cold climatisation. The Tm 12.86
 CC antibody/antisera which is also used as a screening device to screen
 CC cDNA libraries in an expression system, including cross-species cDNA
 CC libraries to identify homologous sequences in other species. The
 CC present sequence represents the Tm13.17 anti-freeze protein of the
 CC invention.

SQ Sequence 134 AA;

Query Match 78.4%; Score 29; DB 23; Length 134;
 Best Local Similarity 62.5%; Pred. No. 1.9e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 RETESTV 8
 :||:|
 Db 109 RDTVEETV 116

RESULT 14
 AAU10049
 ID AAU10049 standard; Protein; 149 AA.
 AC AAU10049;
 XX
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Tm 13.17 clone/His tag fusion protein minus signal sequence.
 XX
 XX Anti-freeze peptide; Tm 13.17; yellow mealworm beetle; THP; cryoinjury;
 KW thermal hysteresis protein; cryosurgery; cold climatisation;
 KW hypothermic cell preservation; de-icing formulation; transgenic plant;
 KW transgenic animal; His tag.
 XX
 OS Tenebrio molitor.
 XX
 FH Key Location/Qualifiers
 FT Region 1..33
 FT /note= "Peptide including His tag"
 FT Protein 34..149
 FT /note= "Mature AFP protein, the signal sequence
 FT has been removed for the purposes of cloning"
 XX
 XX WO200194378-A1.
 XX

PD 13-DEC-2001.
 XX
 PF 07-JUN-2001; 2001WO-US18532.
 XX
 PR 08-JUN-2000; 2000US-210446P.
 XX
 XX (UYNV) UNIV NEW YORK STATE RES FOUND.
 PA (HORN/) HORWATH K L.
 PA (MYER/) MYERS K L.
 PA (EAST/) EASTON C M.
 XX
 XX Horwath KL, Myere KL, Easton CM;
 PI
 DR WPI: 2002-090137/12.
 DR N-PSDB; AAS21120.
 XX
 XX New CDNA polynucleotide encoding a thermal hysteresis protein which is
 PT a Type III anti-freeze protein derived from the Tenebrionidae
 PT Superfamily, useful for providing antifreeze protection to improve the
 PT quality of food -
 XX
 PS Claim 15; Fig 5.12; 364pp; English.
 XX
 CC This invention relates to a CDNA polynucleotide comprising a nucleotide
 CC sequence encoding a thermal hysteresis protein (THP) which is a Type III
 CC anti-freeze protein derived from the Tenebrionidae Superfamily. Anti-
 CC freeze proteins lower the freezing point of a solution without affecting
 CC the melting point of the solution. An activated anti-freeze protein may
 CC be incorporated into plant, produce or fish in an amount sufficient to
 CC provide antifreeze protection or in a region of a target tissue to
 CC provide antifreeze protein to limit tumour cell or target tissue
 CC cryoinjury during cryosurgery. The proteins of the invention may also be
 CC used in hypothermic solutions or bathing media to reduce cold damage in
 CC order to provide cryogenic or hypothermic preservation of cells and
 CC tissues. The proteins may be used as de-icing formulations or used on
 CC surfaces to reduce existing ice buildup or abate the formation of ice
 CC buildup on surfaces such as a road, aircraft, household products,
 CC machinery and plant surfaces or as a food product to improve the quality
 CC of food by abating freezing of solutions, freezer burn, or degradation
 CC due to cold storage. The polynucleotides for the activated protein can
 CC be used to create transgenic or gene-modified plants, crops, fish, or
 CC animals having greater tolerance to cold climatization. The Tm 12.86
 CC antibody/antisera which is also used as a screening device to screen
 CC CDNA libraries in an expression system, including cross-species CDNA
 CC libraries to identify homologous sequences in other species. The
 CC present sequence represents the Tm 13.17 clone/his tag fusion protein
 CC minus the signal peptide of the invention. This CDNA was created to try
 CC to enhance the anti-freeze activity of the recombinant protein.
 CC
 XX
 SQ Sequence 149 AA;
 Query Match 78.4%; Score 29; DB 23; Length 149;
 Best Local Similarity 62.5%; Pred. No. 2.1e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RETESTV 8
 Db 124 RDTVEETV 131
 RESULT 15
 AAU76232
 ID AAU76232 standard; Protein; 174 AA.
 XX
 AC AAU76232;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Tm 13.17 clone/his tag fusion protein #2.
 XX
 XX Anti-freeze peptide; Tm 13.17; yellow mealworm beetle; THP; cryoinjury;
 KW thermal hysteresis protein; cryosurgery; cold climatization;
 KW hypothermic cell preservation; de-icing formulation; transgenic plant;

KW transgenic animal; His tag.
 XX
 OS Tenebrio molitor.
 XX
 FH Key Location/Qualifiers
 FT Region 1..58
 FT /note= "Peptide including His tag
 FT and signal peptide"
 FT 59..174
 FT Protein /note= "Mature Tm 13.17 protein"
 FT
 PN WO200194378-A1.
 XX
 PD 13-DEC-2001.
 XX
 PF 07-JUN-2001; 2001WO-US18532.
 XX
 PR 08-JUN-2000; 2000US-210446P.
 XX
 XX (UYNV) UNIV NEW YORK STATE RES FOUND.
 PA (HORN/) HORWATH K L.
 PA (MYER/) MYERS K L.
 PA (EAST/) EASTON C M.
 XX
 PI Horwath KL, Myers KL, Easton CM;
 DR WPI: 2002-090137/12.
 XX
 XX New CDNA polynucleotide encoding a thermal hysteresis protein which is
 PT a Type III anti-freeze protein derived from the Tenebrionidae
 PT Superfamily, useful for providing antifreeze protection to improve the
 PT quality of food -
 XX
 PS Disclosure; Page 334; 364pp; English.
 XX
 CC This invention relates to a CDNA polynucleotide comprising a nucleotide
 CC sequence encoding a thermal hysteresis protein (THP) which is a Type III
 CC anti-freeze protein derived from the Tenebrionidae Superfamily. Anti-
 CC freeze proteins lower the freezing point of a solution without affecting
 CC the melting point of the solution. An activated anti-freeze protein may
 CC be incorporated into plant, produce or fish in an amount sufficient to
 CC provide antifreeze protection or in a region of a target tissue to
 CC provide antifreeze protein to limit tumour cell or target tissue
 CC cryoinjury during cryosurgery. The proteins of the invention may also be
 CC used in hypothermic solutions or bathing media to reduce cold damage in
 CC order to provide cryogenic or hypothermic preservation of cells and
 CC tissues. The proteins may be used as de-icing formulations or used on
 CC surfaces to reduce existing ice buildup or abate the formation of ice
 CC buildup on surfaces such as a road, aircraft, household products,
 CC machinery and plant surfaces or as a food product to improve the quality
 CC of food by abating freezing of solutions, freezer burn, or degradation
 CC due to cold storage. The polynucleotides for the activated protein can
 CC be used to create transgenic or gene-modified plants, crops, fish, or
 CC animals having greater tolerance to cold climatization. The Tm 12.86
 CC antibody/antisera which is also used as a screening device to screen
 CC CDNA libraries in an expression system, including cross-species CDNA
 CC libraries to identify homologous sequences in other species. The
 CC present sequence represents the Tm 13.17 clone/his tag fusion protein
 CC of the invention. This CDNA was created to facilitate purification
 CC of the Tm 13.17 protein and to try to enhance the anti-freeze activity of
 CC the recombinant protein.
 CC Note: This sequence differs from the sequence given in Figure 5.11
 CC for the Tm 13.17 clone/his tag fusion protein.
 CC
 XX
 SQ Sequence 174 AA;
 Query Match 78.4%; Score 29; DB 23; Length 174;
 Best Local Similarity 62.5%; Pred. No. 2.5e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RETESTV 8
 Db 149 RDTVEETV 156

Wed Nov 26 13:01:40 2003

us-09-230-111c-11.rag

Page 12

Search completed: November 26, 2003, 12:30:21
Job time : 23.1687 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 26, 2003, 12:22:05 ; Search time 3.85542 Seconds
(without alignments)
97.580 Million cell updates/sec

Title: US-09-230-111c-11

Perfect score: 37

Sequence: 1 RETESTV 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	89.2	3131	1 ESYN_FUSEQ	000869 fusarium eq
2	31	83.8	340	1 SED5_YEAST	001590 saccharomyc
3	30	81.1	193	1 THM4_ARATH	09896 arabidopsis
4	30	81.1	307	1 ARCC_HAANI	048285 halobacteri
5	30	81.1	710	1 EFG_BUCBP	P59451 buchnera ap
6	29	78.4	244	1 DCGI_YEAST	P32460 saccharomyc
7	29	78.4	247	1 ECXI_SULTO	097568 sulfolobus
8	29	78.4	414	1 TTYI_HUMAN	P25490 homo sapien
9	29	78.4	414	1 TTYI_MOUSE	000889 mus musculi
10	29	78.4	472	1 YWIE_CABEL	Q23525 caenorhabdi
11	29	78.4	684	1 ISH1_SCHPO	Q97YX6 schizosacch
12	29	78.4	955	1 KINL_LEICH	P46865 leishmania
13	28	75.7	3830	1 SACS_MOUSE	091168 mus musculi
14	28	75.7	123	1 RSBE_HAUMA	P49402 haloraccula
15	28	75.7	218	1 Y232_SYNY3	Q55705 synecocyst
16	28	75.7	223	1 YB43_YEAST	P38304 saccharomyc
17	28	75.7	226	1 CASE_HUMAN	P05814 homo sapien
18	28	75.7	293	1 SCRK_STRMU	Q07211 streptococc
19	28	75.7	322	1 VANI_ENTFC	005709 enterococcu
20	28	75.7	378	1 DCAW_SCHPO	Q9P773 schizosacch
21	28	75.7	383	1 LA_ABDAL	Q26457 aedes albop
22	28	75.7	400	1 TKBI_STRPU	Q26648 strongyloce
23	28	75.7	426	1 C1SY_HELPY	P56062 helicobacte
24	28	75.7	493	1 GALT_LACTC	Q98663 lactococcus
25	28	75.7	531	1 YQ42_BACAN	Q9RMO0 bacillus an
26	28	75.7	535	1 HTR1_HAANI	P33741 halobacteri
27	28	75.7	535	1 HTR1_HAISA	P33955 halobacteri
28	28	75.7	556	1 ASNS_SCHPO	F78753 schizosacch
29	28	75.7	622	1 YAE7_YEAST	P39723 saccharomyc
30	28	75.7	660	1 VNCS_PAVPN	P18547 porcine par
31	28	75.7	662	1 VNCS_PAVPK	P52502 porcine par
32	28	75.7	668	1 VNCS_PPV19	P24842 feline panl
33	28	75.7	668	1 VNCS_MEVA	P27438 mink enteri

34	28	75.7	668	1 VNCS_PAVCN	P19292 canine parv
35	28	75.7	668	1 VNCS_PAVL3	P36311 parvovirus
36	28	75.7	672	1 VNCS_MOMIV	P07300 murine minu
37	28	75.7	672	1 VNCS_MOMIV	P03134 murine minu
38	28	75.7	672	1 VNCS_PAVHH	P03133 hamster par
39	28	75.7	691	1 EFG_CAMJE	Q93116 campylobact
40	28	75.7	691	1 EFG_CAMJE	Q93116 campylobact
41	28	75.7	691	1 EFG_HELPJ	Q92K24 helicobacte
42	28	75.7	702	1 EFG_HELPJ	P56002 helicobacte
43	28	75.7	702	1 EFG_HELPJ	Q50565 chidobacte
44	28	75.7	704	1 EFG_YERPE	Q821b3 yersinia pe
45	28	75.7	708	1 EFGI_STRCO	Q8kag9 chlorobium
					P40173 streptomyce

ALIGNMENTS

RESULT 1
ID ESYN_FUSEQ STANDARD; PRT; 3131 AA.
AC 000869;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Enniatin synthetase (includes: N-methylcyclopeptide synthetase
DE (EC 6.3.2.-); S-adenosyl-L-methionine-dependent N-methyltransferase
DE (EC 2.1.1.-)).
GN ESYN.
OS Fusarium equiseti (Fusarium sclirp).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; mitosporic Hypocreales; Fusarium.
OX NCBI_TaxId=61235;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Lambotte et Fautrey;
RX MEDLINE=93247491; PubMed=8463420;
RA Haese A., Schubert M., Herrmann M., Zocher R.;
RT "Molecular characterization of the enniatin synthetase gene encoding a
RT multifunctional enzyme catalysing N-methylpeptide formation in
RT Fusarium sclirp.";
RL Mol. Microbiol. 7:905-914 (1993).
RN [2]
RP REVISIONS.
RA Zocher R.;
RT Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE OF 1011-1034; 1677-1695; 2029-2049; 2098-2106 AND 2294-2299,
RP AND FUNCTION.
RC STRAIN=Lambotte et Fautrey;
RX MEDLINE=95324513; PubMed=7601090;
RA Pieper R., Haese A., Schroeder W., Zocher R.;
RT "Arrangement of catalytic sites in the multifunctional enzyme enniatin
RT synthetase.";
RL Eur. J. Biochem. 230:119-126 (1995).
RN [4]
RP FUNCTION, AND MUTAGENESIS OF TYR-2106.
RX MEDLINE=20469430; PubMed=10887181;
RA Hacker C., Glinke M., Hornbogen T., Doller A., Zocher R.;
RT "Mutation analysis of the N-methyltransferase domain of the
RT multifunctional enzyme enniatin synthetase.";
RL J. Biol. Chem. 275:30826-30832 (2000).
RN [5]
RP ENZYME REGULATION.
RC STRAIN=Lambotte et Fautrey;
RA Billich A., Zocher R.;
RT "N-methyltransferase function of the multifunctional enzyme enniatin
RT synthetase.";
RL Biochemistry 26:8417-8423 (1987).
CC -1- FUNCTION: A multifunctional enzyme which activates and then
CC thioesterifies adenylated D-2-hydroxyisovaleric acid and a
CC hydrophobic N-methyl-L-amino acid. Peptide bond formation and N-
CC methylation of the amino acid occur before three enzyme-bound
CC dipeptides are condensed to a hexapeptide. This then cyclizes

[illegible]

RESULT 2	SEDS_YEAST	STANDARD;	PRT;	340 AA.
ID	SEDS_YEAST			
AC	Q01590;			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Integral membrane protein SEDS.			
GN	SEDS OR YLR026C.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.			
OX	NCBI_TaxId=4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93016264; PubMed=1400588;			
RT	Hardwick K.G., Pelham H.R.B.;			
RT	"SDS encodes a 39-kD integral membrane protein required for			
RL	vesicular transport between the ER and the Golgi complex.";			
RL	J. Cell Biol. 119:513-521(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	STRAIN=S288C / AB972;			
RX	MEDLINE=97313267; PubMed=9169871;			
RA	Johnston M., Hillier L., Riles L., Albertmann K., Andre B., Anorge W.,			
RA	Benes V., Bruechner M., Delius H., Dubois E., Duesterhoeft A.,			
RA	Enrian K.-D., Floeth M., Goffeau A., Hebling U., Henmann K.,			
RA	Heuss-Netzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,			
RA	Louis E.U., Messengier F., Mewes H.-W., Miosga T., Moestl D.,			
RA	Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,			
RA	Portellelle D., Purnelle B., Reckmann S., Rieger M., Rinke M., Rose M.,			
RA	Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,			
RA	Vierewood A.P., Urrutiarazu L.A., Vandenbol M., Verhaesselt P.,			
RA	Vieredeels F., Voet M., Volckaert G., Voss H., Wambolt R., Wedler E.,			
RA	Weller H., Zimmerman F.K., Zollner A., Hani J., Hobels J.D.;			
RT	"The nucleotide sequence of Saccharomyces cerevisiae chromosome XII";			
RL	Nature 387:87-90(1997).			
CC	-1- FUNCTION: Required for vesicular transport between the			
CC	endoplasmic reticulum and the Golgi complex. Acts as a target			
CC	organellar soluble NSF attachment protein receptor (t-SNARE).			
CC	-1- SUBUNIT: Interacts with STPL and SPB3.			
CC	-1- SUBCELLULAR LOCATION: Type IV membrane protein (Potential).			
CC	-1- SIMILARITY: BELONGS TO THE SYNTAXIN/EPIMORPHIN FAMILY.			
CC	-1- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X66980; CAA47390.1; -			
DR	EMBL; Z73198; CAA97549.1; -			
DR	PIR; A44019; A44019.			
DR	PDB; 1MQS; 20-NOV-02.			
DR	SGD; S0004016; SEDS.			
DR	GO; GO:0005801; C:Golgi cis-face; IDA.			
DR	InterPro; IPR006012; Syntaxin.			
DR	InterPro; IPR006011; Syntaxin_N.			
DR	InterPro; IPR000727; t-SNARE.			
DR	Flam; PF00804; Syntaxin; 1.			
DR	SMART; SMO0397; t-SNARE; 1.			
DR	PROSITE; PS00914; SYNTAXIN; 1.			
DR	PROSITE; PS50192; t-SNARE; 1.			
KW	Coiled coil; Transport; Protein transport; Transmembrane; Golgi stack;			
KW	3D-structure.			
FT	DOMAIN 1 319			
FT	TRANSMEM 320 340			
FT				

```

FT DOMAIN 249 311 T-SNARE COILED-COIL HOMOMOLOGY.
FT DOMAIN 146 173 COILED COIL (POTENTIAL).
SQ SEQUENCE 340 AA; 38807 MW; B3192B57269A0B83 CRC64;

Query Match
Best Local Similarity 83.8%; Score 31; DB 1; Length 340;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RETESTV 8
Db 260 ETEESTI 266

RESULT 3
THM4_ARATH STANDARD; PRT; 193 AA.
AC Q9SEU6; Q9LDP6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thioredoxin M-type-4, chloroplast precursor (TRX-M4).
GN AT3G15360 OR MTK13.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20047910; PubMed=10580150;
RA Meesters-Ortega D., Meyer Y.;
RT "The Arabidopsis thaliana genome encodes at least four thioredoxins m
RT and a new prokaryotic-like thioredoxin.";
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=CV, Columbia;
RC MEDLINE=20277480; PubMed=10819329;
RX Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
RA "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.";
RN [3]
RA DNA Ref. 7.131-135 (2000).
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN=CV, Columbia;
RC Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGEC)."
RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Brewer V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldman K.;
RT "Full-length cDNA from Arabidopsis thaliana."
RT Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Participates in various redox reactions through the
CC reversible oxidation of the active center dihydro to a disulfide.
CC The M form is known to activate NADP-malate dehydrogenase (by
CC similarity).
CC -1- SUBCELLULAR LOCATION: Chloroplast (by similarity).
CC -1- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY. PLANT M-TYPE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, AF095752; AAF15951.1; -
CC
CC EMBL, AB022216; BAB02365.1; -

```

```

DR EMBL, AC024081; AAF35402.1; -
DR EMBL, AF375443; AAK3027.1; -
DR EMBL, AY060538; AAL31169.1; -
DR EMBL, AY088157; AAM65701.1; -
DR HSSP; P80579; 1QW.
DR InterPro; IPR006662; Thior.
DR InterPro; IPR006663; Thior.
DR InterPro; IPR005746; Thior.
DR Pfam; PF00085; Thior.
DR PRINTS; PR00421; THIOREDOXIN.
DR TIGRFAMs; TIGR01068; Thior.
DR PROSITE; PS00194; THIOREDOXIN; 1.
KW Redox-active center; Electron transport; Chloroplast; Transit peptide;
KW Multigene family.
FT TRANSIT 1 82 CHLOROPLAST (POTENTIAL).
FT CHAIN 83 193 THIOREDOXIN M-TYPE 4.
FT DISULFID 116 119 REDOX-ACTIVE (BY SIMILARITY).
FT CONFLICT 151 151 A -> P (IN REF. 1).
SQ SEQUENCE 193 AA; 21172 MW; 7189545252C60DD9D CRC64;

Query Match
Best Local Similarity 81.1%; Score 30; DB 1; Length 193;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RETESTV 8
Db 180 RETLEKTI 187

RESULT 4
ARCC HALNT1 STANDARD; PRT; 307 AA.
ID ARCC HALNT1
AC Q48295; Q9HEM9;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carbamate kinase (EC 2.7.2.2).
GN ARCC OR VNC6316G.
OS Halobacterium sp. (Strain NRC-1 / ATCC 700922 / JCM 11081), and
OS Halobacterium salinarum.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OC NCBI_TaxId=64091, 2242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1; PLASMID=pNRC200;
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahatza G.G., Bergquist B., Pan M.,
RA Shukla H.D., Laskey S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welci R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angewine C.M., Dale H.,
RA Isebnarger T.A., Peck R.F., Pohlroeder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=H. salinarum; STRAIN=L33;
RX MEDLINE=96326340; PubMed=8759859;
RA Ruepp A., Soppe J.;
RT "Fermentative arginine degradation in Halobacterium salinarum
RT (formerly Halobacterium halobium): genes, gene products, and
RT transcripts of the arcrAB gene cluster."
RL J. Bacteriol. 178:4942-4947 (1996).
CC -1- CATALYTIC ACTIVITY: ATP + NH(3) + CO(2) = ADP + carbamoyl
CC phosphate.
CC -1- PATHWAY: Arginine degradation via arginine deiminase; third step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC -1- SIMILARITY: Belongs to the carbamate kinase family.
CC

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005161; CAAG20947.1; -
DR EMBL; X80931; CAA56905.1; -
DR PIR; T44864; T44864.
DR HSSP; P95474; IE19.
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR003964; Bac_carb_kinase.
DR Pfam; PF00696; aak_kinase.1.
DR PRINTS; PR01469; CARMKINASE.
DR TIGRPFAM; TIGR00746; arcc; 1.
KW transferase; Kinase; Arginine metabolism; Plasmid; Complete proteome.
SQ SEQUENCE 307 AA; 32534 MW; B8E58C81ED7B8435 CRC64;

Query Match
Best Local Similarity 81.1%; Score 30; DB 1; Length 307;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RETESTV 8
|:|:|:|
Db 25 RDTIEGT 32

RESULT 5
EFG_BUCBP STANDARD; PRT; 710 AA.
ID EFG_BUCBP
AC P59451;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Elongation factor G (EF-G).
GN FUS1 OR BBP470.
OS Buchnera aphidicola (subsp. Baizongia pistaciae).
OC Bacterii; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=135842;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426901; PubMed=12522265;
RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
RA Bascola U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
RA Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;
RT "Reductive genome evolution in Buchnera aphidicola.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
CC -!- FUNCTION: This protein promotes the GTP-dependent translocation of
CC the nascent protein chain from the A-site to the P-site of the
CC ribosome.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE014017; AAO27176.1; -
DR HAMAP; MF 00054; -; 1.
DR PROSITE; PS00301; EFACITOR GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding;
KW Complete proteome.
FT NP_BIND 17 24 GTP (BY SIMILARITY).
FT NP_BIND 88 92 GTP (BY SIMILARITY).
FT NP_BIND 142 145 GTP (BY SIMILARITY).

SQ SEQUENCE 710 AA; 79363 MW; 037BA37B190557FD CRC64;

Query Match
Best Local Similarity 81.1%; Score 30; DB 1; Length 710;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETESTV 8
|:|:|:|
Db 497 RETIOSSV 504

RESULT 6
DCG1_YEAST STANDARD; PRT; 244 AA.
ID DCG1_YEAST
AC P32460;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE DCG1 protein.
GN DCG1 OR YIR030C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92009196; PubMed=1916277;
RA Yeo H.S., Cooper T.G.;
RT "Sequences of two adjacent genes, one (DGL2) encoding allantoicase
RT and another (DCG1) sensitive to nitrogen-catabolite repression in
RT Saccharomyces cerevisiae.";
RL Gene 104:55-62(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX PubMed=9169870;
RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.E., Hornell T., Hunt S., Jagels K., Jones M., Lye G.,
RA Moulis S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Bartell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
RL Nature 387:84-87(1997).
CC -!- MISCELLANEOUS: SENSITIVE TO NITROGEN-CATABOLITE REPRESSION.
CC -!- SIMILARITY: TO S.POMBE SPAC17.10 AND PSEUDOMONAS HYDANTOIN
CC RACEMASE (HYDE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M64719; AAA34562.1; -
DR EMBL; Z38061; CAA86190.1; -
DR PIR; S48492; S48492.
DR SGD; S0001469; DCG1.
DR GO; GO:0006807; P:nitrogen metabolism; IEP.
SQ SEQUENCE 244 AA; 27313 MW; 8B5261745C0B494 CRC64;

Query Match
Best Local Similarity 78.4%; Score 29; DB 1; Length 244;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RETESTV 7
|:|:|:|
Db 21 RETIEKT 27

RESULT 7
ECK1_SULTO

ID BCX1_SULTO STANDARD; PRT; 247 AA.
AC 097568;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
GN Probable exosome complex exonuclease 1 (EC 3.1.13.-).
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfobacterales; Sulfobacterales;
OC Sulfobacterales.
OX NCBI_TaxID=111955;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawabata Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankaei A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka A., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
RA Aoki K.-I., Maeda S., Yanagita M., Nishimura M., Yamagishi A.,
RA Ohnaka T., Kikuchi H.,
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfobacterales tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
CC -1- FUNCTION: Probably involved in the 3'-5' degradation of a variety
CC of RNA species (Potential).
CC -1- SUBUNIT: Component of the archaeal exosome multienzyme
CC ribonuclease complex (Potential).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO THE RNASE PH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF000982; BAB5433.1; -.
CC DR HAMAP; MF_00591; -1.
CC DR InterPro; IPR001247; 3 EXORNAse.
CC DR Pfam; PF01138; RNase PH; 1.
CC DR Pfam; PF03725; RNase PH; 1.
CC KM Exosome; Hydrolyase; Nuclease; Exonuclease; Complete proteome.
CC SQ SEQUENCE 247 AA; 27644 MW; 7EB042B784C6599 CRC64;
CC -----
CC Query Match 78.4%; Score 29; DB 1; Length 247;
CC Best Local Similarity 62.5%; Prcd. No. 42;
CC Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RA Shi Y., Seto E., Chang L.-S., Shenk T.;
RT "Transcriptional repression by Y1, a human GLI-Kruppel-related
RT protein, and relief of repression by adenovirus E1A protein.";
RL Cell 67:377-388(1991).
RN (12)
RP SEQUENCE FROM N.A.
RC TISSUE=Forebrain;
RX MEDLINE=92052179; PubMed=1946405;
RA Park K., Atchison M.;
RT "Isolation of a candidate repressor/activator, NF-E1 (YY-1, delta),
RT that binds to the immunoglobulin kappa 3' enhancer and the
RT immunoglobulin heavy-chain mu E1 site.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9804-9808(1991).
RN (3)
RP SEQUENCE FROM N.A.
RA Whitson R.H., Huang T., Dang J., Itakura K.;
RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
RN (4)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strassberg R.L., Fehlgold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Mollath S.J.,
RA Bose S.A., McKernan P.J., McKernan K.J., Malik J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Balesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywicki M.T., Skalska U., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN (15)
RP SUBCELLULAR LOCATION:
RX MEDLINE=98152931; PubMed=9493912;
RA McNeil S., Guo B., Stein J.L., Lian J.B., Bushmeyer S., Seto E.,
RA Atchison M.L., Penman S., van Wijnen A.J., Stein G.S.;
RT "Targeting of the YY1 transcription factor to the nucleus and the
RT nuclear matrix in situ: the C-terminus is a principal determinant for
RT nuclear trafficking.";
RL J. Cell. Biochem. 68:500-510(1998).
RN (6)
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 304-414.
RX MEDLINE=97098436; PubMed=8942976;
RA Houdavay H.B., Ubeva A., Shenk T., Burley S.K.;
RT "Crystal structure of YY1 bound to the adenovirus-associated virus P5
RT initiator.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:13577-13582(1996).
RN (7)
RP STRUCTURE BY NMR OF 353-379.
RX MEDLINE=98308000; PubMed=9642075;
RA Viles J.H., Patel S.U., Mitchell J.B.O., Moody C.M., Justice D.E.,
RA Upendrabindu J., Doyle P.W., Harris C.D., Sadler P.J., Thornton J.M.,
RT "Design, synthesis and structure of a zinc finger with an artificial
RT beta-turn.";
RL J. Mol. Biol. 279:973-986(1998).
CC -1- FUNCTION: MULTIFUNCTIONAL TRANSCRIPTION FACTOR THAT EXHIBITS
CC POSITIVE AND NEGATIVE CONTROL ON A LARGE NUMBER OF CELLULAR AND
CC VIRAL GENES BY BINDING TO SITES OVERLAPPING THE TRANSCRIPTION
CC START SITE. MAY PLAY AN IMPORTANT ROLE IN DEVELOPMENT AND
CC DIFFERENTIATION. THE FUNCTION OF YY1 AS AN ACTIVATOR OR A
CC REPRESSOR IS SPECIFIED BY THE PRESENCE OF OTHER PROTEINS. FOR
CC EXAMPLE IT ACTS AS A REPRESSOR IN ABSENCE OF ADENOVIRUS E1A
CC PROTEIN BUT AS AN ACTIVATOR IN ITS PRESENCE.

CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; W76541; AAA5926.1; -;
 CC EMBL; Z14077; CA78455.1; -;
 CC EMBL; EC037308; AAH37308.1; -;
 CC PIR; A40350; A40350.
 CC PDB; 1UBD; 23-DEC-96.
 CC PDB; 1ZNM; 01-APR-98.
 CC TRANSFAC; T00915; -;
 CC DR Gene; HGNC:12856; Y1.
 CC MIM; 600013; -;
 CC DR GO; GO:0003713; F:transcription co-activator activity; TAS.
 CC DR GO; GO:0003714; F:transcription co-repressor activity; TAS.
 CC DR GO; GO:0003700; F:transcription factor activity; TAS.
 CC DR GO; GO:0008270; F:zinc ion binding activity; TAS.
 CC DR GO; GO:0006960; P:antimicrobial humoral response (sensu Inver. . .); TAS.
 CC DR GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; TAS.
 CC DR InterPro; IPR007087; ZnF_C2H2.
 CC DR Pfam; PF00096; zf-C2H2; 4.
 CC DR ProDom; PD000003; ZnF_C2H2; 1.
 CC DR SMART; SM00355; ZnF_C2H2; 4.
 CC DR PROSITE; PS00028; ZINC_FINGER_C2H2_2; 4.
 CC DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.
 CC Transcription regulation; Repressor; Activator; Nuclear protein;
 CC Zinc-finger; Metal-binding; DNA-binding; Repeat; 3D-structure.
 CC KW DOMAIN 43 53
 CC FT DOMAIN 54 69
 CC FT DOMAIN 70 80
 CC FT DOMAIN 159 170
 CC FT ZN_FING 296 320
 CC FT ZN_FING 325 347
 CC FT ZN_FING 353 377
 CC FT ZN_FING 383 407
 CC FT DOMAIN 257 341
 CC FT DOMAIN 333 371
 CC FT DOMAIN 371 397
 CC FT CONFLICT 65 65
 CC FT CONFLICT 196 196
 CC FT STRAND 296 297
 CC FT TURN 301 302
 CC FT STRAND 306 307
 CC FT HELIX 310 317
 CC FT HELIX 318 320
 CC FT STRAND 325 326
 CC FT STRAND 328 330
 CC FT STRAND 333 334
 CC FT HELIX 337 343
 CC FT HELIX 344 346
 CC FT TURN 347 347
 CC FT STRAND 353 354
 CC FT TURN 358 359
 CC FT STRAND 363 364
 CC FT HELIX 367 378
 CC FT TURN 388 389
 CC FT TURN 397 400
 CC FT HELIX 401 407
 CC SEQUENCE 414 AA; 44712 MW; 058C05A0AD2D04E6 CRC64;
 Query Match 78.4%; Score 29; DB 1; Length 414;
 Best local similarity 85.7%; Pred. No. 73;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ETESTV 8
 Db 34 ETESTV 40
 RESULT 9
 ID TY11_MOUSE STANDARD; PRT; 414 AA.
 AC Q00899;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Transcriptional repressor protein YY1 (Yin and Yang 1) (YY-1) (Delta
 DE transcription factor) (NF-E1) (UCR-motif DNA-binding protein).
 GN YY1 OR UCRBP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92107191; PubMed=1309593;
 RA Flanagan J.R., Becker K.G., Enlist D.L., Gleason S.L., Driggers P.H.,
 RA Levi B.-Z., Appella E., Ozato K.;
 RT "Cloning of a negative transcription factor that binds to the
 RT upstream conserved region of Moloney murine leukemia virus.";
 RL Mol. Cell. Biol. 12:38-44 (1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=93296177; PubMed=8516301;
 RA Safary G., Perry R.P.;
 RT "Characterization of the mouse gene that encodes the delta/YY1/NF-
 RT E1/UCRB transcription factor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:5559-5563 (1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92052178; PubMed=1946404;
 RA Hartharan N., Kelley D.E., Perry R.P.;
 RT "Delta, a transcription factor that binds to downstream elements in
 RT several polymerase II promoters, is a functionally versatile zinc
 RT finger protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9799-9803 (1991).
 CC -1- FUNCTION: MULTIFUNCTIONAL TRANSCRIPTION FACTOR THAT EXHIBITS
 CC POSITIVE AND NEGATIVE CONTROL ON A LARGE NUMBER OF CELLULAR AND
 CC VIRAL GENES BY BINDING TO SITES OVERLAPPING THE TRANSCRIPTION
 CC START SITE. MAY PLAY AN IMPORTANT ROLE IN DEVELOPMENT AND
 CC DIFFERENTIATION. THE FUNCTION OF YY1 AS AN ACTIVATOR OR A
 CC REPRESSOR IS SPECIFIED BY THE PRESENCE OF OTHER PROTEINS. BINDS TO
 CC THE UPSTREAM CONSERVED REGION (UCR) (5'-GCCATTTT-3') OF MOLONEY
 CC MURINE LEUKEMIA VIRUS (MULV).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; W73963; AAA40522.1; -;
 CC EMBL; L13968; AAA40477.1; -;
 CC EMBL; L13969; AAA40477.1; JOINED.
 CC EMBL; L13965; AAA40477.1; JOINED.
 CC EMBL; L13966; AAA40477.1; JOINED.
 CC EMBL; L13967; AAA40477.1; JOINED.
 CC EMBL; W74590; AAA37521.1; -;
 CC PIR; A48273; A48273.

```

DR HSP; P25490; IUBD.
DR TRANSFAC; T00278; -.
DR TRANSFAC; T00865; -.
DR MGI; 99150; Y1L.
DR GO; GO:0005667; C:transcription factor complex; IDA.
DR GO; GO:0003700; P:transcription factor activity; IDA.
DR GO; GO:0006355; P:regulation of transcription; IDA.
DR InterPro; IPR007087; Znf.C2H2.
DR Pfam; PF00086; Zf-C2H2; 1.
DR ProDom; PD000003; Znf.C2H2; 1.
DR SMART; SM00355; Znf.C2H2; 4.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 4.
KW Transcription regulation; Repressor; Activator; Nuclear protein;
KW Zinc-finger; Metal-binding; DNA-binding; Repeat.
FT DOMAIN 43 53
FT DOMAIN 54 70
FT DOMAIN 71 82
FT DOMAIN 161 170
FT ZN_PING 161 170
FT ZN_PING 296 320
FT ZN_PING 325 347
FT ZN_PING 353 377
FT ZN_PING 383 407
FT DOMAIN 257 341
FT DOMAIN 333 371
FT DOMAIN 371 397
FT DOMAIN 371 397
FT CONFLICT 219 219
FT CONFLICT 375 375
FT CONFLICT 414 AA; 44717 MW; C012378285984F9 CRC64;
SQ SEQUENCE 414 AA; 44717 MW; C012378285984F9 CRC64;

Query Match 78.4%; Score 29; DB 1; Length 414;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RETEST 8
DB 34 RETEST 40

RESULT 10
YMW_CAEEL STANDARD; PRT; 472 AA.
AC Q23525;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 54.3 kDa protein ZK546.14 in chromosome II.
GN ZK546.14.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peldoderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Briceol N2;
RA Hallsworth K.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SURF6 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U29380; AAA68745.1; -.
CC PIR; T27903; T27903.

```

```

DR WormPep; ZK546.14; CE02914.
DR InterPro; IPR007019; SURF6.
DR Pfam; PF04935; SURF6; 1.
KW Hypothetical protein; Nuclear protein.
SQ SEQUENCE 472 AA; 54261 MW; 56EA30387545D3D CRC64;

Query Match 78.4%; Score 29; DB 1; Length 472;
Best Local Similarity 71.4%; Pred. No. 84;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RETEST 7
DB 302 RETEST 308

RESULT 11
ID ISH1_SCHPO STANDARD; PRT; 684 AA.
AC Q9YXG6; Q09149;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Stress response protein Ish1.
DE ISH1 OR ISP1 OR SPB365.12C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_Taxid=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAP1;
RA Turt T.G., Mueller U.W., Sazer S., Rose J.K.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
[2]
SQ SEQUENCE FROM N.A.
RC STRAIN=9721-.
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felkewell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle S.J., Hunt S., Jasele K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitz E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkhardt G., Aert R., Robben J., Grynolprez B.,
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabell C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leinrich H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wandut R., Purnelle B.,
RA Colfeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Nottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garçon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potaashkin J.,
RA Shpakovski G.V., Ueberry D., Barrett B.G., Nurse P.;
RA The genome sequence of Schizosaccharomyces pombe.
RL Nature 415:871-880(2002).
[3]
CC FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
RP MEDLINE=21895766; PubMed=11751918;
RX Taricant L., Tejada M.L., Young P.G.;
RT "The fission yeast E52 homologue, Bist, interacts with the Ish1
RT stress-responsive nuclear envelope protein";
RL J. Biol. Chem. 277:10562-10572(2002).
CC -1- FUNCTION: Has a role in maintaining cell viability during
CC stationary phase induced by stress response. Activated by the spc1

```

CC MAPK pathway.
 CC -1- SUBUNIT: Homodimer, and heterodimer with bist.
 CC -1- SUBCELLULAR LOCATION: Nuclear envelope.
 CC -1- SIMILARITY: BELONGS TO THE LEA TYPE 1 FAMILY.
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 606.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U38783; AAC9262.1; ALT_FRAME.
 CC EMBL: AL078627; CAB44764.1; -.
 CC PIR: T40319; T40319.
 CC GeneDB: SPombe; SPBC365.12c; -.
 CC InterPro: IPR004238; LEA.
 CC Pfam: PF02987; LEA. 2.
 CC NUCLEAR PROTEIN.
 CC CONFLICT 418 418 S -> C (IN REF. 1).
 CC CONFLICT 490 490 V -> E (IN REF. 1).
 CC SEQUENCE 684 AA; 75698 MW; 884CB69714BC7FAB CRC64;
 SQ
 Query Match 78.4%; Score 29; DB 1; Length 684;
 Best Local Similarity 62.5%; Pred. No. 1.2e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RETESTV 8
 DB 450 RETIDETI 457
 RESULT 12
 KINL LEICH STANDARD; PRT; 955 AA.
 AC P46655;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Kinesin-like protein K39 (Fragment).
 GN KIN.
 OS Leishmania chagasi.
 OC Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatida; Leishmania.
 OX NCBI_TaxID=44271;
 RP SEQUENCE FROM N.A.
 RC STRAIN=MHOM/BR/82 / Isolate BA-2;
 RX MEDLINE=93133867; PubMed=8421715;
 RA Burns J.M. Jr., Shreffler W.G., Benson D.R., Ghallib H.W., Badaro R.,
 RA Reed S.G.;
 RT "Molecular characterization of a kinesin-related antigen of
 RT Leishmania chagasi that detects specific antibody in African and
 RT American visceral leishmaniasis."
 RL Proc. Natl. Acad. Sci. U.S.A. 90:775-779(1993).
 CC -1- DEVELOPMENTAL STAGE: PREDOMINANT IN AMASTIGOTES.
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L07879; AAA29254.1; -.
 CC PIR: A47334; A47334.
 CC HSSP: P17119; 3KAR.
 CC InterPro: IPR001752; kinesin_motor.
 CC Pfam: PF00225; kinesin; 1.

DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC. 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
 KW Motor protein; Microtubules; ATP-binding; Coiled coil; Repeat.
 FT DOMAIN 1 399 KINESIN-MOTOR (BY SIMILARITY).
 FT DOMAIN 426 >955 KINLEID COIL (POTENTIAL).
 FT NP BIND 122 129 COILED COIL (POTENTIAL).
 FT DOMAIN 704 >955 7 X 39 AA APPROXIMATE TANDEM REPEATS.
 FT REPEAT 704 742 1.
 FT REPEAT 743 781 2.
 FT REPEAT 782 820 3.
 FT REPEAT 821 859 4.
 FT REPEAT 860 898 5.
 FT REPEAT 899 937 6.
 FT REPEAT 938 >955 7 (PARTIAL).
 FT NON_TER 955 955
 SQ SEQUENCE 955 AA; 106168 MW; 8CA76815B84C6E9 CRC64;
 QY 1 RETESTV 8
 DB 542 REKLESTV 549
 RESULT 13
 SACS MOUSE STANDARD; PRT; 3830 AA.
 AC O9JUG8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Sacsin.
 GN SACS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20120709; PubMed=10655055;
 RA Engert J.C., Berube P., Mercier J., Dore C., LePage P., Ge B.,
 RA Bouchard J.-P., Mathieu J., Melancon S.B., Schalling M., Lander E.S.,
 RA Morgan K., Hudson T.J., Richter A.;
 RT "ARSACS, a spastic ataxia common in northeastern Quebec, is caused by
 RT mutations in a new gene encoding an 11.5-kb ORF."
 RL Nat. Genet. 24:120-125(2000).
 CC -1- FUNCTION: MAY FUNCTION IN CHAPERONE-MEDIATED PROTEIN FOLDING.
 CC -1- SIMILARITY: Contains 1 U domain.
 CC -1- SIMILARITY: Contains 1 HEPN domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF193557; AAF1263.1; -.
 CC MGI: MGI:1354724; Sacs.
 CC InterPro: IPR001623; DnaJ_N
 CC PROSITE; PS00636; DnaJ_1; FALSE_NEG.
 CC PROSITE; PS50076; DnaJ_2; 1.
 DR PROSITE; PS50910; HEPN; 1.
 KW Chaperone.
 FT DOMAIN 3557 3644 J-DOMAIN.
 FT DOMAIN 3702 3818 HEPN.
 SQ SEQUENCE 3830 AA; 436750 MW; B3F9DA5A3B0C3EPI CRC64;

Query Match 78.4%; Score 29; DB 1; Length 3830;
 Best Local Similarity 62.5%; Pred. No. 7.7e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 RETESTV 8
 Db 2868 KETLOSTV 2875

RESULT 14
 RS9E HALMA STANDARD; PRT; 123 AA.
 ID RS9E HALMA
 AC P49402;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 308 ribosomal protein S8e (HS23).
 GN RPS8E.
 OS Halococcus marismortui (Halobacterium marismortui).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 CC Halobacteriaceae; Halococci.
 RX NCBI_TaxID=2238;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE=95391073; PubMed=7662106;
 RA Engemann S., Herluth E., Briesemanster U., Wittmann-Liebold B.,
 RT "Amino acid sequence of the ribosomal protein HS23 from the
 RT halophilic Halococcus marismortui and homology studies to other
 RT ribosomal proteins."
 RL J. Protein Chem. 14:189-195(1995).
 CC -1- MASS SPECTROMETRY: MW=13552.4; MW-ERR=13; METHOD=Electrospray.
 CC -1- SIMILARITY: BELONGS TO THE S8E FAMILY OF RIBOSOMAL PROTEINS.
 DR PIR; S49022; S49022.
 DR HAMAP; MP_00029; -; 1.
 DR InterPro; IPR001047; Ribosomal_S8E.
 DR Pfam; PF01201; Ribosomal_S8e; 1.
 DR PRODOM; PD005658; Ribosomal_S8E; 1.
 DR TIGRPFAM; TIGR00307; S8e; 1.
 DR PROSITE; PS01193; RIBOSOMAL_S8E; 1.
 KW Ribosomal protein.
 SQ SEQUENCE 123 AA; 13549 MW; FB7DB07AC3B1D0AD CRC64;

Query Match 75.7%; Score 28; DB 1; Length 123;
 Best Local Similarity 71.4%; Pred. No. 34;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 ETTESTV 8
 Db 68 ETTESTI 74

RESULT 15
 Y232 SYNY3 STANDARD; PRT; 218 AA.
 ID Y232 SYNY3
 AC 055705;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein slr0232.
 GN slr0232.
 OS Synechocystis sp. (Strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 CC NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96127529; PubMed=8590279;
 RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
 RA Sugitara M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
 RT region from map positions 644 to 924 of the genome."
 RL DNA Res. 2:153-166(1995).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: BELONGS TO THE DEDA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; D64000; BAA10237.1; -;
 DR PIR; S76385; S76385.
 DR InterPro; IPR000352; Deda.
 DR Pfam; PF00597; Deda; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 218 AA; 23781 MW; C04B5D7B7EA7F863 CRC64;

Query Match 75.7%; Score 28; DB 1; Length 218;
 Best Local Similarity 71.4%; Pred. No. 61;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 ETTESTV 8
 Db 207 QTTESTI 213

Search completed: November 26, 2003, 12:31:15
 Job time: 5.85542 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:25:40 ; Search time 7.42169 Seconds
(without alignments)
103.662 Million cell updates/sec

Title: US-09-230-111c-11

Perfect score: 37

Sequence: 1 RETESTV 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	89.2	3131	2 S39842	emulatin synthetase
2	32	86.5	281	2 A69099	fumarate hydratase
3	31	83.8	340	2 A44019	membrane protein S
4	30	81.1	69	2 A10473	hypothetical prote
5	30	81.1	307	2 T44864	carbamate kinase (
6	30	81.1	632	2 T06586	DNA-binding protei
7	30	81.1	648	2 AF3181	hypothetical prote
8	30	81.1	650	2 T04830	probable serine/th
9	30	81.1	894	2 T27007	hypothetical prote
10	29	78.4	89	2 T36743	hypothetical prote
11	29	78.4	244	2 S48492	DCG1 protein - yea
12	29	78.4	267	2 A84233	hypothetical prote
13	29	78.4	268	2 C81325	probable periplasm
14	29	78.4	373	2 JC2426	transcription acti
15	29	78.4	389	2 AF0920	uroporphyrinogen I
16	29	78.4	414	2 A40350	transcription repr
17	29	78.4	414	2 A48273	delta-VYL/NF-EL/UC
18	29	78.4	435	2 T33199	hypothetical prote
19	29	78.4	441	2 E72242	Mg-protoporphyryn
20	29	78.4	468	2 B83046	probable class III
21	29	78.4	472	2 T27903	hypothetical prote
22	29	78.4	488	2 T71943	1-aminocyclopropan
23	29	78.4	684	2 T40319	hypothetical prote
24	29	78.4	740	2 T26569	hypothetical prote
25	29	78.4	955	2 A47334	lectin kinesin-rela
26	29	78.4	1132	2 T45037	hypothetical prote
27	29	78.4	1139	2 B70954	hypothetical prote
28	28	75.7	123	1 S49022	ribosomal protein
29	28	75.7	137	2 G70215	hypothetical prote

30	28	75.7	208	2 T41555	hypothetical prote
31	28	75.7	218	2 S76385	hypothetical prote
32	28	75.7	223	2 S34018	hypothetical prote
33	28	75.7	226	1 KBHU	beta-casein precu
34	28	75.7	255	2 T44991	oxidoreductase [m
35	28	75.7	322	2 JQ1191	vanh protein - Ent
36	28	75.7	323	2 T17660	hypothetical prote
37	28	75.7	333	2 T19314	hypothetical prote
38	28	75.7	379	2 AC1767	UDP-N-acetylglucos
39	28	75.7	379	2 A11391	UDP-N-acetylglucos
40	28	75.7	395	2 C90108	phocoystem II sta
41	28	75.7	426	2 B64523	citrate synthase -
42	28	75.7	430	2 A97518	dihydroorotase (dh
43	28	75.7	430	2 AC2737	dihydroorotase [m
44	28	75.7	440	2 T34621	NADH2 dehydrogenas
45	28	75.7	441	2 G82612	hypothetical prote

ALIGNMENTS

RESULT 1

S39842
emulatin synthetase - fungus (Fusarium sclrtpl)
C:Species: Fusarium sclrtpl
C>Date: 31-Dec-1993 #sequence revision 02-Jun-1994 #ext_change 03-Nov-2000
C/Accession: S39842; S35906; S65363
R:Haese, A.
submitted to the EMBL Data Library, November 1992
A:Reference number: S39842
A:Accession: S39842

A:Molecule type: DNA
A:Residues: 1-3131 <HAE>
A:Cross-references: EMBL:Z18755; NID:92729; PIDN:CAA79245.1; PID:92730
R:Haese, A.; Schubert, M.; Herrmann, M.; Zocher, R.
Mol. Microbiol. 7, 905-914, 1993
A>Title: Molecular characterization of the emulatin synthetase gene encoding a multifun
A:Reference number: S35906; MUID:93247491; PMID:8483420
A:Accession: S35906
A>Status: nucleic acid sequence not shown

A:Molecule type: DNA
A:Residues: 499-1074;1572-1988;2423-2566 <HA2>
A:Cross-references: EMBL:Z18755
A:Experimental source: strain ETH 1536/J5
R:Pieper, R.; Haese, A.; Schroeder, W.; Zocher, R.
Eur. J. Biochem. 230, 119-126, 1995
A>Title: Arrangement of catalytic sites in the multifunctional enzyme emulatin synthetase
A:Reference number: S65363; MUID:95324513; PMID:7601090
A:Accession: S65363

A:Molecule type: protein
A:Residues: 2029-2048;430-437;1011-1020;1021-1034;1677-1695;2294-2299 <PIE>
A:Experimental source: strain ETH 1536/J5
C:Genetics:
A:Gene: esyl1
C:Superfamily: acetate-CoA ligase homology; acyl carrier protein homology
C:Keywords: carrier protein; multifunctional enzyme; phosphopantetheine; phosphoprotein
F:531-988/Domain: acetate-CoA ligase homology <ACLI1>
F:1603-2100/Domain: acetate-CoA ligase homology <ACLI2>
F:1507-2574/Domain: acyl carrier protein homology <ACP1>
F:12601-2667/Domain: acyl carrier protein homology <ACP2>
F:1047,2538,2633/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 89.2%; Score 33; DB 2; Length 3131;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OV 1 RETESTV 8
DB 593 RETVOSTV 600

RESULT 2

A69099

fumarate hydratase, class I - Methanobacterium thermoautotrophicum (strain Delta H)
C/Species: Methanobacterium thermoautotrophicum
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 16-Jul-1999
C/Accession: A69099
R/Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadford, R.; Vitell, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A/Reference number: A69000; MUID:98037514; PMID:9371463
A/Accession: A69099
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-281 <MTH>
A/Cross-references: GB:AE000929; GB:AE000666; NID:g2622853; PIDN:AB86205.1; PID:g262286
A/Experimental source: strain Delta H
C/Genetics:
A/Gene: MTH1735
A/Suprafamily: Iron-dependent tarrate dehydratase alpha chain; iron-dependent tarrate
F:52-249/Domain: iron-dependent tarrate dehydratase alpha chain homology <TTDA>

Query Match 86.5%; Score 32; DB 2; Length 281;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETESTV 8
DB 4 RETESTV 11

RESULT 3
A44019
membrane protein SED5 - yeast (Saccharomyces cerevisiae)
C/Species: Saccharomyces cerevisiae
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jul-2000
C/Accession: A44019; S64853; S23223
R/Hardwick, K.G.; Pelham, H.R.B.
J. Cell Biol. 119, 513-521, 1992
A/Title: SED5 encodes a 35-kD integral membrane protein required for vesicular transport
A/Reference number: A44019; MUID:93016284; PMID:1400588
A/Accession: A44019
A/Molecule type: DNA
A/Residues: 1-340 <HAR>
A/Cross-references: EMBL:X66980; NID:g4455; PIDN:CAA47390.1; PID:g4456
A/Note: sequence extracted from NCBI backbone (NCBIR:116626)
R/Obermayer, B.; Pitravandi, E.; Rinke, M.
submitted to the Protein Sequence Database, May 1996
A/Reference number: S64845
A/Accession: S64853
A/Molecule type: DNA
A/Residues: 1-340 <OBE>
A/Cross-references: EMBL:Z73198; NID:g1360335; PID:e245501; PID:g1360336; MIPS:YL026C
A/Note: experimental_source strain S288C
C/Genetics:
A/Gene: SGD:SED5
A/Cross-references: SGD:S0004016; MIPS:YL026C
A/Map position: 12R
C/Keywords: transmembrane protein
F:323-339/Domain: transmembrane #status predicted <TM>

Query Match 83.8%; Score 31; DB 2; Length 340;
Best Local Similarity 85.7%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ETTESTV 8
DB 260 ETTESTV 266

RESULT 4
A10473
hypocretin-like protein YPO3890 [imported] - Yersinia pestis (strain CO92)

C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C/Accession: A10473
R/Parhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Parraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; MUID:21470413; PMID:11586360
A/Accession: A10473
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-69 <YPO>
A/Cross-references: GB:AL590842; PIDN:CA093357.1; PID:g15981803; GSPDB:GN00175
C/Genetics:
A/Gene: YPO3890

Query Match 81.1%; Score 30; DB 2; Length 69;
Best Local Similarity 75.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETESTV 8
DB 9 KETIEGTV 16

RESULT 5
T44864
cardamate kinase (EC 2.7.2.2) [imported] - Halobacterium salinarum
C/Species: Halobacterium salinarum
C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C/Accession: T44864
R/Ruopp, A.; Soppa, J.
J. Bacteriol. 178, 4942-4947, 1996
A/Title: Fermentative arginine degradation in Halobacterium salinarum (formerly Halobac
A/Reference number: Z22862; MUID:96326340; PMID:8759859
A/Accession: T44864
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-307 <RUB>
A/Cross-references: EMBL:X80931; NID:g984738; PIDN:CAA56905.1; PID:g984741
A/Experimental source: strain L33
C/Genetics:
A/Gene: arcC
C/Suprafamily: cardamate kinase
C/Keywords: phosphotransferase

Query Match 81.1%; Score 30; DB 2; Length 307;
Best Local Similarity 75.0%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETESTV 8
DB 25 RDTIEQTV 32

RESULT 6
T06586
DNA-binding protein PD2 - garden pea
C/Species: Pisum sativum (garden pea)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 08-Oct-1999
C/Accession: T06586
R/Sato, N.; Kazuno, A.A.; Ohra, N.; Ohshima, K.
submitted to the EMBL Data Library, June 1996
A/Description: Identification of a novel family of DNA-binding proteins with two AT-hook
A/Reference number: Z15774
A/Accession: T06586
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-632 <SAT>
A/Cross-references: EMBL:X98740; NID:e995229; PIDN:CAA67292.1; PID:e275185
A/Experimental source: cv. Alaska

Query Match 81.1%; Score 30; DB 2; Length 632;
 Best Local Similarity 62.5%; Pred. No. 1.5e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 RETESTV 8
 :|||:|
 Db 158 KETVENTV 165

RESULT 7

AF3181

hypothetical protein Atu5178 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C:Species: Agrobacterium tumefaciens
 C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C:Accession: AF3181
 R:Wood, D.W.; Seubald, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erege, G.; Gallet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McClellan
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Woo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 eter, E.W.

A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; WUID:21608550; PMID:11743193

A:Accession: AF3181

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-648 <KUR>

A:Cross-references: GB:AE008667; PIDN:AAL45868.1; PID:G17743610; GSPDB:GN00188

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu5178

C:Superfamily: Escherichia coli probable ABC transporter yjbZ, ATP-binding cassette hom

Query Match 81.1%; Score 30; DB 2; Length 648;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RETEST 7
 :|||:|
 Db 511 RETEST 517

RESULT 8

T04830

probable serine/threonine-specific protein kinase (EC 2.7.1.-) F21P8.20 - Arabidopsis th
 N:Alternate names: protein F21P8.20; protein F7H19.320

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 31-Jan-2000

C:Accession: T04830; T05145

R:Bevan, M.; Massenet, O.; Clabault, G.; Outgley, F.; Maché, R.; Bancroft, I.; Mewes, H.
 submitted to the Protein Sequence Database, August 1998

A:Reference number: Z15386

A:Accession: T04830

A:Molecule type: DNA

A:Residues: 1-650 <BEV>

A:Cross-references: EMBL:AL022347

A:Experimental source: cultivar Columbia; BAC clone F21P8

R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew
 submitted to the Protein Sequence Database, July 1998

A:Reference number: Z15399

A:Accession: T05145

A:Molecule type: DNA

A:Residues: 1-650 <BEW>

A:Cross-references: EMBL:AL031018

A:Experimental source: cultivar Columbia; BAC clone F7H19

C:Genetics:

A:Map position: 4

A:introns: 112/1; 303/1; 347/3; 418/1; 497/2; 548/3

A>Note: F21P8.20; F7H19.320

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase, homolo
 C:Keywords: phosphotransferase; protein kinase

Query Match 81.1%; Score 30; DB 2; Length 650;
 Best Local Similarity 75.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETESTV 8
 :|||:|
 Db 228 KETIBECTV 235

RESULT 9

T27007

hypothetical protein Y48B6A.11 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27007
 R:Wall, M.
 submitted to the EMBL Data Library, September 1999

A:Reference number: Z20297
 A:Accession: T27007

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-894 <WIL>

A:Cross-references: EMBL:AL110490; NID:e1542263; PIDN:CAB54451.1; CESP:Y48B6A.11

A:Experimental source: clone Y48B6A

C:Genetics:

A:Gene: CESP:Y48B6A.11

A:introns: 71/2; 146/2; 510/1; 558/2

Query Match 81.1%; Score 30; DB 2; Length 894;
 Best Local Similarity 85.7%; Pred. No. 2.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RETEST 7
 :|||:|
 Db 149 RETEST 155

RESULT 10

T36743

hypothetical protein SC11.10c - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T36743
 R:Saunders, D.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, July 1999

A:Reference number: Z21613

A:Accession: T36743

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-89 <SNU>

A:Cross-references: EMBL:AL096849; PIDN:CAB50936.1; GSPDB:GN00070; SCOEDB:SC11.10c

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC11.10c

Query Match 78.4%; Score 29; DB 2; Length 89;
 Best Local Similarity 71.4%; Pred. No. 2;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RETEST 7
 :|||:|
 Db 41 RETEST 47

RESULT 11

S48492

DG1 protein - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YIR030c
 C:Species: Saccharomyces cerevisiae

C>Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 18-Feb-2000
 C:Accession: S48492; S19038

R:Rowley, K.
 submitted to the EMBL Data Library, October 1994

A:Reference number: S48478
A:Accession: S48492
A:Molecule type: DNA
A:Residues: 1-244 <NON>
A:Cross-references: GB:Z47047; EMBL:Z38061; NID:g603997; PID:g763375; MIPS:YIR030C
R:Yoo, H.S.; Cooper, T.G.
Gene 104, 55-62, 1991
A:Title: Sequences of two adjacent genes, one (DNL2) encoding allantoinase and another (DNL1) encoding allantoinase.
A:Reference number: JH0442; MUID:92009196; PMID:1916277
A:Accession: S19038
A:Molecule type: DNA
A:Residues: 1-126, 'C', 128-244 <YOO>
A:Cross-references: GB:M64719
C:Genetics:
A:Gene: SGD.DCG1
A:Cross-references: SGD:S0001469; MIPS:YIR030C
A:Map position: 9R
C:Superfamily: Saccharomyces cerevisiae DCG1 protein
C:Keywords: transmembrane protein
F:221-237/Domain: transmembrane #status predicted <TMM>

Query Match 78.4%; Score 29; DB 2; Length 244;
Best Local Similarity 85.7%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETEST 7
DB 21 RETEST 27

RESULT 12
A64233
Hypothetical protein Vng0754c [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: A64233
R:Ng, W.V.; Kennedy, S.P.; Malaitas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.; Leihanser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabcic Jung, K.H.; Alam, M.; Freltas, T.; 12176-12181, 2000
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A64160; MUID:20504483; PMID:11016950
A:Accession: A64233
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-267 <STO>
A:Cross-references: GB:A6004437; NID:G10580333; PID:AA619229.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNG0754C

Query Match 78.4%; Score 29; DB 2; Length 267;
Best Local Similarity 62.5%; Pred. No. 96;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 RETEST 8
DB 166 KESVESTV 173

RESULT 13
C81325
Probable periplasmic protein Cj1193c [imported] - Campylobacter jejuni (strain NCTC 1116 C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: C81325
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Basham, D.; Chilling C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanYlist, A.; Whitehead, S.; Barrell Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: C81325
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-268 <PAR>
A:Cross-references: GB:AL139077; GB:AL111168; NID:g6968444; PID:CA673447.1; PID:g696862 A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj1193c

Query Match 78.4%; Score 29; DB 2; Length 268;
Best Local Similarity 62.5%; Pred. No. 96;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETEST 8
DB 224 RETVDDTV 231

RESULT 14
JC2426
transcription activator/repressor protein delta/yy1 - African clawed frog
N:Alternate names: gene F11 protein
C:Species: Xenopus laevis (African clawed frog)
C:Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 21-Jul-2000
C:Accession: JC2426; S42062
R:Pisaneschi, G.; Cecchetti, S.; Falchetti, M.L.; Finicino, S.; Carnevali, F.; Beccari, Biochem. Biophys. Res. Commun. 205, 1236-1242, 1994
A:Title: Characterization of F11/YY1, a Xenopus laevis conserved zinc-finger protein b A:Reference number: JC2426; MUID:95100950; PMID:7802655
A:Accession: JC2426
A:Molecule type: mRNA
A:Residues: 1-373 <PIS>
A:Cross-references: EMBL:X77698; NID:g455607; PID:CAA54777.1; PID:g455608 A:Note: The authors translated the codon TTG for residue 71 as Val and GCG for residue 1 C:Keywords: zinc finger
F:257-366/Region: zinc fingers

Query Match 78.4%; Score 29; DB 2; Length 373;
Best Local Similarity 85.7%; Pred. No. 14e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RETEST 8
DB 36 RETESTV 42

RESULT 15
AF0920
urotophyrinogen III methylase [imported] - Salmonella enterica subsp. enterica serovar uroporphyrinogen III methylase
C:Species: Salmonella enterica subsp. enterica serovar Typh
A:Note: this species has also been called Salmonella typh
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AF0920
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, th, T.; Connerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moutle, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K. A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AF0920
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-389 <PAR>
A:Cross-references: GB:AL13382; PID:CAD09384.1; PID:g16504502; GSPDB:GN00176
C:Genetics:
A:Gene: STY3623

Query Match 78.4%; Score 29; DB 2; Length 389;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RETEST 7
DB 15 RETEST 21

Wed Nov 26 13:01:40 2003

us-09-230-111c-11.rpr

Page 5

Search completed: November 26, 2003, 12:35:59
Job time : 9.42169 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:25:10 ; Search time 17.4458 Seconds

(Without alignments)
118.334 Million cell updates/sec

Title: US-09-230-111c-11

Perfect score: 37
Sequence: 1 RETIESTV 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp_undefined:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	91.9	1026	5 Q8MMU3	Q8MMU3 dicystostellin
2	32	86.5	281	17 Q27768	Q27768 mechanobact
3	30	81.1	69	16 Q8ZACO	Q8ZACO yersinia pe
4	30	81.1	108	16 Q8KY52	Q8KY52 ralsinonla pe
5	30	81.1	275	17 Q8FT27	Q8FT27 mechanobact
6	30	81.1	289	16 Q8NSVO	Q8NSVO corynebacte
7	30	81.1	433	16 Q8F8R6	Q8F8R6 leptospira
8	30	81.1	632	10 Q04697	Q04697 plasmu activ
9	30	81.1	648	16 Q8UKX4	Q8UKX4 agrobacteri
10	30	81.1	650	10 Q54465	Q54465 arabidopsis
11	30	81.1	659	10 Q9C588	Q9C588 arabidopsis
12	30	81.1	663	10 Q8H1S3	Q8H1S3 arabidopsis
13	30	81.1	713	16 Q8F3H8	Q8F3H8 leptospira
14	30	81.1	894	5 Q9U297	Q9U297 caenorhabd
15	30	81.1	1051	10 Q8R2Y7	Q8R2Y7 oryza sativ
16	29	76.4	89	16 Q9S2A9	Q9S2A9 streptomyce

17	29	78.4	144	16 Q8DKS1	Q8DKS1 synechococc
18	29	78.4	200	12 Q91LK3	Q91LK3 white spot
19	29	78.4	247	17 Q975G8	Q975G8 sulfoblobus
20	29	78.4	258	5 Q8MYL9	Q8MYL9 caenorhabd
21	29	78.4	267	17 Q9HRC9	Q9HRC9 halobacteri
22	29	78.4	268	16 Q9PNA2	Q9PNA2 campylobact
23	29	78.4	278	2 Q93BE5	Q93BE5 mycobacteri
24	29	78.4	300	11 Q8C6B5	Q8C6B5 mus musculu
25	29	78.4	373	13 Q91745	Q91745 xenopus lae
26	29	78.4	389	16 Q8Z399	Q8Z399 salmoneilla
27	29	78.4	392	12 Q9DMA6	Q9DMA6 rat cytoleg
28	29	78.4	411	11 Q8CHY4	Q8CHY4 rattus norv
29	29	78.4	422	3 Q8X009	Q8X009 neurospora
30	29	78.4	435	5 Q61867	Q61867 caenorhabd
31	29	78.4	441	16 Q9X1M3	Q9X1M3 thermotoga
32	29	78.4	468	16 Q9HV04	Q9HV04 pseudomonas
33	29	78.4	484	11 Q8BSFO	Q8BSFO mus musculu
34	29	78.4	488	10 Q43368	Q43368 arabidopsis
35	29	78.4	541	10 Q8LL16	Q8LL16 chlamydomon
36	29	78.4	733	4 Q9H1W1	Q9H1W1 homo sapien
37	29	78.4	740	5 Q9U2P7	Q9U2P7 caenorhabd
38	29	78.4	897	5 Q9BHM4	Q9BHM4 pleistophor
39	29	78.4	1096	11 Q8BKW6	Q8BKW6 mus musculu
40	29	78.4	1147	16 Q50466	Q50466 mycobacteri
41	29	78.4	1151	4 Q9Y2H6	Q9Y2H6 homo sapien
42	29	78.4	1182	10 Q9FT71	Q9FT71 arabidopsis
43	29	78.4	1188	10 Q8L840	Q8L840 arabidopsis
44	29	78.4	1341	12 Q998F6	Q998F6 porcine cyt
45	29	78.4	5060	2 Q9K5M1	Q9K5M1 anabaena sp

ALIGNMENTS

RESULT 1

ID Q8MMU3 PRELIMINARY; PRT; 1026 AA.

AC Q8MMU3;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE GTPase-activator protein for RHO-like GTPases.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szefranek K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox B., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium."
RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AC117073; AAM43650.1;
DR InterPro; IPR000198; RhogAP.
DR Pfam; PF00620; RhogAP; 1.
DR SMART; SM00324; RhogAP; 1.
SQ SEQUENCE 1026 AA; 116407 MW; 47CC4400F76D893A CRC64;

Query Match
Best Local Similarity 91.9%; Score 34; DB 5; Length 1026;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RETIESTV 8

DB 861 RETIESTL 866

RESULT 2

ID Q27768 PRELIMINARY; PRT; 281 AA.
AC Q27768;
DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
 DE Fumarate hydratase, class I.
 GN MTH1735.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Delta H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
 RA Spadofora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Sater H., Patwell D., Prabhakar S.,
 RA McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.,
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genomics."
 RL J. Bacteriol. 179:7135-7155(1997).
 DR EMBL; AE00929; AAB86205.1; -;
 DR InterPro; IPR004646; Ttda_fumA_fumb;
 DR TIGRFAMs; TIGR00722; ttda_fumA_fumb; 1.
 KW Complete proteome.
 SQ SEQUENCE 281 AA; 30252 MW; 6A0C03B0F1908976 CRC64;

Query Match 86.5%; Score 32; DB 17; Length 281;
 Best Local Similarity 75.0%; Pred. No. 55;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETESTV 8
 DB 4 RETESTV 11

RESULT 3
 O8ZAC0 PRELIMINARY; PRT; 69 AA.
 AC O8ZAC0;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)
 DE Hypothetical protein YP03890.
 GN YP03890.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parthall J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sepahia M., James K.D., Churcher C., Hengall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Felwell T., Hamlin N., Holroyd S., Jagsels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.,
 RT "Genome sequence of Yersinia pestis, the causative agent of plague."
 RT Nature 413:523-527(2001).
 DR EMBL; AJ414159; CAC93357.1; -;
 DR Hypothetical protein; Complete proteome.
 KW SEQUENCE 69 AA; 7118 MW; FB1BA82011D6D2C7 CRC64;

Query Match 81.1%; Score 30; DB 16; Length 69;
 Best Local Similarity 75.0%; Pred. No. 34;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETESTV 8

DB 9 KETIGTV 16

RESULT 4
 O8XYS2 PRELIMINARY; PRT; 108 AA.
 AC O8XYS2;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
 DE Probable phage HK022 Gp9-related protein.
 GN RSC1686 OR RS04080.
 OC Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Ralstoniaceae; Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GMI100;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Arlat M., Billault A., Brottier P., Camus J.C., Catolico L.,
 RA Chander M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,
 RA Gaspin C., Lavié M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Signier P., Trebault P., Whalen M., Wincker P., Levy M.,
 RA Weisenbach J., Boucher C.A.,
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
 RL Nature 415:497-502(2002).
 DR EMBL; AL646065; CAD15388.1; -;
 DR InterPro; IPR006453; Gp16_SPP1.
 DR TIGRFAMs; TIGR01563; gp16_SPP1; 1.
 KW Complete proteome.
 SQ SEQUENCE 108 AA; 12298 MW; 2D5E38CE1C7E24BC CRC64;

Query Match 81.1%; Score 30; DB 16; Length 108;
 Best Local Similarity 75.0%; Pred. No. 55;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RETESTV 8
 DB 52 RETESTV 59

RESULT 5
 O8PT27 PRELIMINARY; PRT; 275 AA.
 ID O8PT27;
 AC O8PT27;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
 DE ABC transporter, permease protein.
 GN MM2891.
 OS Methanosarcina mazei (Methanosarcina frisia).
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2209;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Goe1 / Goe1 / ATCC BAA-199 / DSM 3647 / OCM 88;
 RX MEDLINE=22120827; PubMed=12125824;
 RA Dependier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
 RA Martinez-Arias R., Henne A., Wietzer A., Baumer S., Jacobi C.,
 RA Bruggemann H., Lienard T., Christmann A., Boencke M., Steckel S.,
 RA Bhattacharya A., Lykide A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
 RA Fritz H.-U., Gottschalk G.,
 RT "The genome of Methanosarcina mazei: evidence for lateral gene
 RT transfer between Bacteria and Archaea."
 RT J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
 DR EMBL; AE013540; AAM32587.1; -;
 DR InterPro; IPR003439; ABC_transporter.
 DR Pfam; PF02687; DUF214; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.

KW Complete Proteome.
SQ SEQUENCE 275 AA; 29798 MW; C43CB22EDE7137F2 CRC64;

Query Match 81.1%; Score 30; DB 17; Length 275;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETEST 7
DB 130 KETEST 136

RESULT 6

Q8NSV0 PRELIMINARY; PRT; 289 AA.
AC Q8NSV0;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DE Pseudouridylylate synthase (EC 4.2.1.70).
CN CGU0567.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005275; BAB97960.1; Pseudou synth_1.
DR InterPro; IPR001406; Pseudou synth_1.
DR Pfam; PF01416; Pseudou synth_1; 2.
DR TIGRFAMs; TIGR00071; hist_true_1.
KW Lyase; Complete Proteome.
SQ SEQUENCE 289 AA; 31992 MW; BC0D2876B61D6E2C CRC64;

Query Match 81.1%; Score 30; DB 16; Length 289;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RETEST 8
DB 41 RETEST 48

RESULT 7

Q8F8R6 PRELIMINARY; PRT; 433 AA.
AC Q8F8R6;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE ATP-dependent DNA helicase pcrh (EC 3.6.1.-).
CN PCRh1 OR LA0468.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=171;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar 1a1;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB011235; AAN47687.1; -.
KW Helicase; Hydrolyase; Complete Proteome.
SQ SEQUENCE 433 AA; 50579 MW; 1CB9364C5C17573F CRC64;

Query Match 81.1%; Score 30; DB 16; Length 433;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETEST 7

DB 410 RETEST 416

RESULT 8

O04697 PRELIMINARY; PRT; 632 AA.
AC O04697;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DE DNA-binding protein PD2.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Alaska; TISSUE=leaf;
RA Sato N., Kazuno A.A., Ohta N., Ohshima K.;
RT "Identification of a novel family of DNA-binding proteins with two AT-hook motifs from pea."
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X98740; CAA67292.1; -.
KW DNA-binding.
SQ SEQUENCE 632 AA; 69498 MW; 9F744E27CD08717 CRC64;

Query Match 81.1%; Score 30; DB 10; Length 632;
Best Local Similarity 62.5%; Pred. No. 3.8e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETEST 8
DB 158 KETEST 165

RESULT 9

Q8UXE4 PRELIMINARY; PRT; 648 AA.
AC Q8UXE4;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE ABC transporter, nucleotide binding/ATPase protein.
GN ATU5178 OR AGR_PAT_247.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
[1]
RN SEQUENCE FROM N.A.
RC MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D., Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-Y., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Neeter E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens C58."
RL Science 294:2317-2323 (2001).
[2]
RN SEQUENCE FROM N.A.
RC MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Ouriello B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,
RA Houniel K., Gordon J., Vaudin M., Tatchouk O., Epp A., Liu F.,

RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58."
 RL Science 294:2323-2328 (2001).
 DR EMBL; AE008940; AAL45868.1; -;
 DR EMBL; AE007888; AAK90549.1; -;
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_transporter.
 DR InterPro; IPR003838; DUF214.
 DR Pfam; PF00005; ABC_tran; 1.
 DR Pfam; PF02687; DUF214; 1.
 DR ProDom; PD000006; ABC_transporter; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KM PfamId; Complete proteome.
 SQ SEQUENCE 648 AA; 69173 MW; 7DBFC05156AD7E47 CRC64;

Query Match 81.1%; Score 30; DB 16; Length 648;
 Best Local Similarity 85.7%; Pred. No. 3.9e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RETEST 7
 Db 511 RETIETT 517

RESULT 10
 ID 065465 PRELIMINARY; PRT; 650 AA.

AC 065465;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Protein kinase - like protein (Protein kinase-like protein).
 GN P21P8.20 OR F7H19.320 OR A1G23130.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Masenget O., Clabault G., Quigley F., Mache R., Bancroft I.,
 RA Mewes H.W., Mayer K., Schueller C.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Peters S.A., van Staveren M., Dirkse W., Stiekema W., Bancroft I.,
 RA Mewes H.W., Mayer K.F.X., Schueller C., Bevan M.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,
 RA Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Masenget O., Clabault G., Quigley F., Mache R., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AL022347; CAI18460.1; -;
 DR EMBL; AL031018; CAI19829.1; -;
 DR EMBL; AL161558; CAB79268.1; -;
 DR InterPro; IPR002902; DUF26.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.

DR Pfam; PF01657; DUF26; 2.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 650 AA; 72246 MW; 40C8AC3B87C5F629 CRC64;

Query Match 81.1%; Score 30; DB 10; Length 650;
 Best Local Similarity 75.0%; Pred. No. 3.9e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RETESTV 8
 Db 228 KETIEGIV 235

RESULT 11
 ID 09C5S8 PRELIMINARY; PRT; 659 AA.

AC 09C5S8;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Receptor-like protein kinase 6.
 GN RXK6
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20575726; PubMed=11135117;
 RA Du L., Chen Z.;
 RT "Identification of genes encoding receptor-like protein kinases as
 RT possible targets of pathogen- and salicylic acid-induced WRKY DNA-
 RT binding proteins in Arabidopsis."
 RL Plant J. 24:837-847(2000). THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AF224707; AAK8317.1; -;
 DR InterPro; IPR002902; DUF26.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF01657; DUF26; 2.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 659 AA; 73402 MW; 5FE8E6A0FBC3D96 CRC64;

Query Match 81.1%; Score 30; DB 10; Length 659;
 Best Local Similarity 75.0%; Pred. No. 3.9e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RETESTV 8
 Db 245 KETIEGIV 252

RESULT 12
 ID 08H1S3 PRELIMINARY; PRT; 663 AA.

AC 08H1S3;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein.


```

GN AT4G23130.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucoside II; Brassicales; Brassicaceae; Arabidopsie.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Huan V.W.,
RA Lee J.M., Quach H.L., Tang C.C., Teriumi M., Wallender E.K., Wong C.,
RA Wu H.C., Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J.,
RA Nguyen M., Palm C.J., Shinn P., Southwick A., Tripp M.G., Wu T.,
RA Davis R.W., Ecker J.R., Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF142496; AF13047.1; -.
KW Hypothetical protein.
SQ
Query Match
Best Local Similarity 81.1%; Score 30; DB 10; Length 663;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RETESTV 8
Db 245 RETESTV 252

RESULT 13
OQ8F3H8 PRELIMINARY; PRT; 713 AA.
ID OQ8F3H8;
AC OQ8F3H8;
DT 01-MAR-2003 (TRENBLREL. 23, Created)
DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)
DE 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
DE Chemotaxis protein Chea.
GN CHEA2 OR LA2425.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxId=173;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar 1a1;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB011410; AA049624.1; -.
KW Complete proteome.
SQ SEQUENCE 713 AA; 79974 MW; A8B1F55842F4BDA CRC64;

Query Match
Best Local Similarity 81.1%; Score 30; DB 16; Length 713;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RETESTV 8
Db 590 RETESTV 597

RESULT 14
OQ9U297 PRELIMINARY; PRT; 894 AA.
ID OQ9U297;
AC OQ9U297;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
DE Y48B6A.11 protein.
GN Y48B6A.11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peleodierinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA Wall M.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=99069613; PubMed=9851916;
RX none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AL110490; CAB54451.1; -.
DR WormBep; Y48B6A.11; CE22126.
DR InterPro; IPR000345; Cyto_heme_bind.
DR InterPro; IPR003347; TF_JmjC.
DR InterPro; IPR003349; TF_JmjC.
DR InterPro; IPR001965; ZnF_PHD.
DR Pfam; PF02373; JmjC; 1.
DR Pfam; PF02375; JmjC; 1.
DR SMART; SM00558; JmjC; 1.
DR SMART; SM00545; JmjC; 1.
DR SMART; SM00243; PHD; 2.
DR PROSITE; PS00190; CYTOCHROME_C; 2.
SQ SEQUENCE 894 AA; 102118 MW; 922FACDEFDD61C47 CRC64;

Query Match
Best Local Similarity 81.1%; Score 30; DB 5; Length 894;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETESTV 7
Db 149 RETESTV 155

RESULT 15
OQ8RZV7 PRELIMINARY; PRT; 1051 AA.
ID OQ8RZV7;
AC OQ8RZV7;
DT 01-JUN-2002 (TRENBLREL. 21, Created)
DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)
DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
DE Myosin heavy chain-like protein.
GN P0034C09.25.
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhacoidae; Oryzaceae; Oryza.
OX NCBI_TaxId=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Nipponbare;
RA Sasaki T., Matsunoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, PAC
RT clone: P0034C09.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003450; BAB84612.1; -.
DR Gramene; OQ8RZV7; -.
SQ SEQUENCE 1051 AA; 117272 MW; 90C1B4F2830D83AD CRC64;

Query Match
Best Local Similarity 81.1%; Score 30; DB 10; Length 1051;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETESTV 8
Db 872 RETESTV 879

```

Search completed: November 26, 2003, 12:34:30
 Job time : 19.4458 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:26:20 ; Search time 19.8795 Seconds
(without alignments)
102.059 Million cell updates/sec

Title: US-09-230-111c-12
Perfect score: 56
Sequence: 1 QNFRYYISFV 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 673684 seqs, 184443283 residues
Total number of hits satisfying chosen parameters: 673684

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	11	8	US-08-681-219-14
2	56	100.0	11	12	US-10-093-138-12
3	56	100.0	11	12	US-09-230-111c-12
4	36	64.3	154	11	US-09-764-891-2398
5	35	62.5	1624	14	US-10-090-454-2
6	35	62.5	1624	15	US-10-090-454-2
7	33	58.9	31	11	US-09-764-891-4866
8	33	58.9	31	15	US-10-205-428-391
9	33	58.9	126	12	US-10-041-860-133
10	33	58.9	126	12	US-10-041-860-208
11	33	58.9	126	12	US-10-041-860-270
12	33	58.9	1543	15	US-10-005-338B-8
13	32	57.1	135	12	US-10-306-878-9
14	32	57.1	186	11	US-09-866-050A-635
15	32	57.1	195	10	US-09-984-245-151

16	32	57.1	195	11	US-09-966-262-151	Sequence 151, App
17	32	57.1	195	11	US-09-983-966-151	Sequence 151, App
18	32	57.1	195	12	US-10-306-878-7	Sequence 7, App1
19	32	57.1	195	15	US-10-143-090-151	Sequence 151, App
20	32	57.1	213	9	US-09-815-242-13993	Sequence 13993, A
21	32	57.1	270	10	US-09-738-626-4836	Sequence 4836, Ap
22	32	57.1	350	15	US-10-112-356-7	Sequence 7, App1
23	32	57.1	350	15	US-10-225-567A-164	Sequence 164, App
24	32	57.1	363	16	US-10-262-313-8	Sequence 8, App1
25	32	57.1	363	16	US-10-262-313-9	Sequence 8, App1
26	32	57.1	366	11	US-09-992-331-8	Sequence 8, App1
27	32	57.1	421	11	US-09-992-331-9	Sequence 9, App1
28	32	57.1	421	11	US-09-759-130B-76	Sequence 76, App1
29	32	57.1	681	11	US-09-759-130B-75	Sequence 75, App1
30	32	57.1	688	12	US-10-032-385-7508	Sequence 7508, Ap
31	32	57.1	700	10	US-09-881-752A-180	Sequence 180, App
32	32	57.1	715	11	US-09-759-130B-73	Sequence 73, App1
33	32	57.1	715	14	US-10-003-132-2	Sequence 2, App1
34	32	57.1	725	12	US-10-032-585-7526	Sequence 7526, Ap
35	32	57.1	1025	12	US-10-354-358-112	Sequence 112, App
36	32	57.1	1844	12	US-10-262-794A-53	Sequence 53, App1
37	32	57.1	1844	15	US-10-242-056-53	Sequence 53, App1
38	32	57.1	2504	9	US-09-817-514A-8	Sequence 8, App1
39	32	57.1	2504	12	US-10-262-794A-12	Sequence 12, App1
40	32	57.1	2504	15	US-10-242-056-12	Sequence 12, App1
41	32	57.1	3353	9	US-09-888-615-64	Sequence 64, App1
42	32	57.1	4351	12	US-09-970-944-20	Sequence 20, App1
43	31	55.4	213	9	US-09-815-242-10170	Sequence 10170, A
44	31	55.4	228	15	US-10-128-714-2331	Sequence 3231, Ap
45	31	55.4	238	10	US-09-925-300-1026	Sequence 1026, Ap

ALIGNMENTS

RESULT 1
US-08-681-219-14
Sequence 14, Application US/08681219
Publication No. US20020058607A1
GENERAL INFORMATION:
APPLICANT: Takaaki Sato and Junn Yanagisawa
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEIN AND THE GLGF
TITLE OF INVENTION: (PDZ/DHR) DOMAIN AND USES THEREOF
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/681,219
FILING DATE: 22-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28, 678
REFERENCE/DOCKET NUMBER: 0575/48962/JPW/JKM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-681-219-14

Query Match 100.0%; Score 56; DB 8; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QNFRTYVSFV 11
DB 1 QNFRTYVSFV 11

RESULT 2

US-10-092-138-12
Sequence 12, Application US/10092138
Publication No. US20030170723A1
GENERAL INFORMATION:
APPLICANT: Sato, Taka-Aki
TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
FILE REFERENCE: 65823/JPM/PT
CURRENT FILING DATE: 2002-09-06
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 12
LENGTH: 11
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial
US-10-092-138-12

Query Match 100.0%; Score 56; DB 12; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QNFRTYVSFV 11
DB 1 QNFRTYVSFV 11

RESULT 3

US-09-230-111C-12
Sequence 12, Application US/09230111C
Publication No. US20030203414A1
GENERAL INFORMATION:
APPLICANT: Sato, Taka-Aki
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN
TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGP (PDZ/DHR)
FILE REFERENCE: 48962-A-PCT-US
CURRENT APPLICATION NUMBER: US/09/230,111C
CURRENT FILING DATE: 1999-05-17
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 12
LENGTH: 11
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial
US-09-230-111C-12

Query Match 100.0%; Score 56; DB 12; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QNFRTYVSFV 11

DB 1 QNFRTYVSFV 11

RESULT 4

US-09-764-891-2998
Sequence 2998, Application US/09764891
Publication No. US20030077808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2998
LENGTH: 154
TYPE: PRT
ORGANISM: Homo sapiens
US-09-764-891-2998

Query Match 64.3%; Score 36; DB 11; Length 154;
Best Local Similarity 55.6%; Pred. No. 34;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 FRTYVSFV 11
DB 105 FRTYVSFV 113

RESULT 5

US-10-090-454-2
Sequence 2, Application US/10090454
Publication No. US20020123106A1
GENERAL INFORMATION:
APPLICANT: Chen, Hongyun
APPLICANT: Le Bihan, Stephane
APPLICANT: Nathani, Parimal S.
APPLICANT: Connop, Bruce P.
TITLE OF INVENTION: NOVEL ABCA9 TRANSPORTER AND USES THEREOF
FILE REFERENCE: 100103.405
CURRENT APPLICATION NUMBER: US/10/090,454
CURRENT FILING DATE: 2002-03-04
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1624
TYPE: PRT
ORGANISM: Homo sapiens
US-10-090-454-2

Query Match 62.5%; Score 35; DB 14; Length 1624;
Best Local Similarity 55.6%; Pred. No. 6.1e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 FRTYVSFV 11
DB 1120 FRTYVSFV 1128

RESULT 6

US-10-005-338B-7
Sequence 7, Application US/10005338B
Publication No. US20030044895A1
GENERAL INFORMATION:
APPLICANT: DENEFLE, Patrice
APPLICANT: ROSIER-MONTUS, Marie-Francoise
APPLICANT: PRADES, Catherine
APPLICANT: ARNOULD-REGUIGNE, Isabelle
APPLICANT: DUVERGER, Nicolas
APPLICANT: ALLIEMETS, Rando

```
/ APPLICANT: DEAN, Michael
/ TITLE OF INVENTION: NUCLEIC ACIDS OF THE HUMAN ABCA5, ABCA6, ABCA9, AND ABCA10 GENES
/ TITLE OF INVENTION: CONTAINING SUCH NUCLEIC ACIDS, AND USES THEREOF
/ FILE REFERENCE: ABCA5, 6, 9, 10
/ CURRENT APPLICATION NUMBER: US/10/005,338B
/ PRIOR FILING DATE: 2001-12-07
/ PRIOR APPLICATION NUMBER: US 60/263,231
/ PRIOR FILING DATE: 2001-01-23
/ PRIOR APPLICATION NUMBER: FR 00403440.1
/ PRIOR FILING DATE: 2000-12-07
/ NUMBER OF SEQ ID NOS: 217
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 7
/ LENGTH: 1624
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-005-338B-7

Query Match      62.5%; Score 35; DB 15; Length 1624;
Best Local Similarity 55.6%; Pred. No. 6.1e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      3 FRTYVSFV 11
DB      1120 FLTIVISF 1128

RESULT 7
US-09-764-891-4866
/ Sequence 4866, Application US/09764891
/ Publication No. US2003007808A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PC006
/ CURRENT APPLICATION NUMBER: US/09/764,891
/ CURRENT FILING DATE: 2001-01-17
/ Prior application data removed - consult PALM or file wrapper
/ NUMBER OF SEQ ID NOS: 10231
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 4866
/ LENGTH: 31
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-764-891-4866

Query Match      58.9%; Score 33; DB 11; Length 31;
Best Local Similarity 55.6%; Pred. No. 23;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      3 FRTYVSFV 11
DB      21 FRTYIVTYL 29

RESULT 8
US-10-205-428-391
/ Sequence 391, Application US/10205428
/ Publication No. US20030108907A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PA17C1
/ CURRENT APPLICATION NUMBER: US/10/205,428
/ CURRENT FILING DATE: 2002-07-26
/ PRIOR APPLICATION NUMBER: 09/764,892
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: 60/179,065
/ PRIOR FILING DATE: 2000-01-31
/ PRIOR APPLICATION NUMBER: 60/180,628
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: 60/214,886
/ PRIOR FILING DATE: 2000-06-28
```

```
/ PRIOR APPLICATION NUMBER: 60/217,487
/ PRIOR FILING DATE: 2000-07-11
/ PRIOR APPLICATION NUMBER: 60/225,758
/ PRIOR FILING DATE: 2000-08-14
/ PRIOR APPLICATION NUMBER: 60/220,963
/ PRIOR FILING DATE: 2000-07-26
/ PRIOR APPLICATION NUMBER: 60/217,496
/ PRIOR FILING DATE: 2000-07-11
/ PRIOR APPLICATION NUMBER: 60/225,447
/ PRIOR FILING DATE: 2000-08-14
/ PRIOR APPLICATION NUMBER: 60/218,290
/ PRIOR FILING DATE: 2000-07-14
/ Remaining prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1019
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 391
/ LENGTH: 31
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-205-428-391

Query Match      58.9%; Score 33; DB 15; Length 31;
Best Local Similarity 55.6%; Pred. No. 23;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      3 FRTYVSFV 11
DB      21 FRTYIVTYL 29

RESULT 9
US-10-041-860-13
/ Sequence 13, Application US/10041860
/ Publication No. US20030157109A1
/ GENERAL INFORMATION:
/ APPLICANT: Corvalan, Jose R.F.
/ APPLICANT: Jia, Xiao-Chi
/ APPLICANT: Feng, Xiao
/ APPLICANT: Yang, Xiao-Dong
/ APPLICANT: Chen, Francine
/ APPLICANT: Gazit, Gadl
/ APPLICANT: Weber, Richard
/ APPLICANT: Bezabeh, Binyam
/ TITLE OF INVENTION: ANTIBODIES DIRECTED TO PDGFR AND USES
/ TITLE OF INVENTION: THEREOF
/ FILE REFERENCE: ABGENIX.051A
/ CURRENT APPLICATION NUMBER: US/10/041,860
/ CURRENT FILING DATE: 2002-01-07
/ NUMBER OF SEQ ID NOS: 377
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 13
/ LENGTH: 126
/ TYPE: PRT
/ ORGANISM: homo sapiens
US-10-041-860-13

Query Match      58.9%; Score 33; DB 12; Length 126;
Best Local Similarity 60.0%; Pred. No. 99;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      2 NERTYVSFV 11
DB      28 NERTYMNWV 37

RESULT 10
US-10-041-860-208
/ Sequence 208, Application US/10041860
/ Publication No. US20030157109A1
/ GENERAL INFORMATION:
/ APPLICANT: Corvalan, Jose R.F.
/ APPLICANT: Jia, Xiao-Chi
/ APPLICANT: Feng, Xiao
```

```
APPLICANT: Yang, Xiao-Dong
APPLICANT: Chen, Francine
APPLICANT: Gazit, Gad
APPLICANT: Weber, Richard
APPLICANT: Bezabeh, Binyam
TITLE OF INVENTION: ANTIBODIES DIRECTED TO PDGFR AND USES
FILE REFERENCE: ABGENIX.051A
CURRENT APPLICATION NUMBER: US/10/041,860
CURRENT FILING DATE: 2002-01-07
NUMBER OF SEQ ID NOS: 377
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 208
LENGTH: 126
TYPE: PRT
ORGANISM: homo sapiens
US-10-041-860-208
```

Query Match 58.9%; Score 33; DB 12; Length 126;
Best Local Similarity 60.0%; Pred. No. 99;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 NRTYIVSFV 11
|||||
Db 28 NRTYIMNV 37

```
RESULT 11
US-10-041-860-270
Sequence 270, Application US/10041860
GENERAL INFORMATION:
APPLICANT: Corvalan, Jose R.F.
APPLICANT: Jia, Xiao-Chi
APPLICANT: Peng, Xiao
APPLICANT: Yang, Xiao-Dong
APPLICANT: Chen, Francine
APPLICANT: Gazit, Gad
APPLICANT: Weber, Richard
APPLICANT: Bezabeh, Binyam
TITLE OF INVENTION: ANTIBODIES DIRECTED TO PDGFR AND USES
FILE REFERENCE: ABGENIX.051A
CURRENT APPLICATION NUMBER: US/10/041,860
CURRENT FILING DATE: 2002-01-07
NUMBER OF SEQ ID NOS: 377
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 270
LENGTH: 126
TYPE: PRT
ORGANISM: homo sapiens
US-10-041-860-270
```

Query Match 58.9%; Score 33; DB 12; Length 126;
Best Local Similarity 60.0%; Pred. No. 99;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 NRTYIVSFV 11
|||||
Db 28 NRTYIMNV 37

```
RESULT 12
US-10-005-338B-8
Sequence 8, Application US/10005338B
Publication No. US20030044895A1
GENERAL INFORMATION:
APPLICANT: DENEUFLE, Patrice
APPLICANT: ROSIER-MONTUS, Marie-Francoise
APPLICANT: PRADES, Catherine
APPLICANT: ARNOULD-REGIGNE, Isabelle
APPLICANT: DUVERGER, Nicolas
APPLICANT: ALLIMETS, Rando
```

```
APPLICANT: DEAN, Michael
TITLE OF INVENTION: NUCLEIC ACIDS OF THE HUMAN ABCA5, ABCA6, ABCA9, AND ABCA10 GENE
TITLE OF INVENTION: CONTAINING SUCH NUCLEIC ACIDS, AND USES THEREOF
FILE REFERENCE: ABCA5, 6, 9, 10
CURRENT APPLICATION NUMBER: US/10/005,338B
CURRENT FILING DATE: 2001-12-07
PRIOR APPLICATION NUMBER: US 60/263,231
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: FR 00403440.1
PRIOR FILING DATE: 2000-12-07
NUMBER OF SEQ ID NOS: 217
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 1543
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: 181
OTHER INFORMATION: Xaa=unknown, may be any amino acid
US-10-005-338B-8
```

Query Match 58.9%; Score 33; DB 15; Length 1543;
Best Local Similarity 55.6%; Pred. No. 1.3e+03;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 PRTYIVSFV 11
|||||
Db 1029 PRTYIVSFV 1037

```
RESULT 13
US-10-306-878-9
Sequence 9, Application US/10306878
Publication No. US20030175819A1
GENERAL INFORMATION:
APPLICANT: Reed, John C.
APPLICANT: Guo, Bin
TITLE OF INVENTION: Methods for Identifying Modulators of
FILE REFERENCE: P-LT 5535
CURRENT APPLICATION NUMBER: US/10/306,878
CURRENT FILING DATE: 2002-11-27
PRIOR APPLICATION NUMBER: US 60/334,149
PRIOR FILING DATE: 2001-11-28
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 135
TYPE: PRT
ORGANISM: Homo sapiens
US-10-306-878-9
```

Query Match 57.1%; Score 32; DB 12; Length 135;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QNRTYIVS 9
|||||
Db 120 QNRTYIVS 128

```
RESULT 14
US-09-866-050A-635
Sequence 635, Application US/09866050A
Publication No. US20030040471A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Muirson, James G.
APPLICANT: Kumbie, Krishanand D.
```

```

; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 635
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-635

Query Match      57.1%; Score 32; DB 11; Length 186;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY      1 QNF--RTYIVSF 10
      :|||:|
Db      100 ENFARTYVSF 111

RESULT 15
US-09-984-245-151
; Sequence 151, Application US/09984245
; Patent NO. US20020165374A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/984,245
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,261
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,168
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160

```

```

; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 151
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-245-151

Query Match      57.1%; Score 32; DB 10; Length 195;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 QNFRTYIVS 9
      :|||:|
Db      180 QNFRTYIVS 188

```

Search completed: November 26, 2003, 12:38:39
 Job time : 20.8795 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using BW model

Run on: November 26, 2003, 12:21:29 ; Search time 30.4819 Seconds
(without alignments)
57.280 Million cell updates/sec

Title: US-09-230-111c-12

Perfect score: 56

Sequence: 1 QNFRYYISFV 11

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	11	19	AAW50177
2	37	66.1	203	23	ABP25867
3	37	66.1	850	24	ABP81504
4	37	66.1	850	24	ABU02762
5	36	64.3	142	21	AA642749
6	36	64.3	142	21	AA655302
7	36	64.3	154	22	AAW94340
8	36	64.3	225	21	AA642748
9	36	64.3	225	21	AA655301

10	36	64.3	232	21	AA642747	Arabidopsis thaliana
11	36	64.3	232	21	AA655300	Arabidopsis thaliana
12	36	64.3	386	22	ABP61754	Drosophila melanog
13	36	64.3	510	23	ABP92124	Herbicidally activ
14	36	64.3	510	24	ABP81234	Arabidopsis thaliana
15	36	64.3	1580	23	ABP81581	Human ABC-A-8-1 pr
16	36	64.3	1581	23	ABP52097	Homo sapiens ABC t
17	35	62.5	592	24	ABU11759	Human MDR1 polypep
18	35	62.5	659	23	ABP49403	Listeria monocytog
19	35	62.5	715	22	ABP63050	Drosophila melanog
20	35	62.5	774	22	ABP71927	Drosophila melanog
21	35	62.5	1467	23	AAU78283	Human apoptosis ac
22	35	62.5	1585	23	AAE21186	Human TRICH-30 pro
23	35	62.5	1624	23	ABP81576	Human ATP-binding
24	35	62.5	1624	24	ABP81582	Human ABC-A-9-3 pr
25	35	62.5	1624	23	ABP84638	Human ABCA9 transp
26	35	62.5	1808	23	ABP92731	Herbicidally activ
27	35	62.5	2291	15	AAW55694	Carbamoyl-phosphat
28	34	60.7	62	22	AAW87832	Human immune/haema
29	34	60.7	90	22	AAW83573	Human immune/haema
30	34	60.7	112	22	AAW09737	Human polypeptide
31	34	60.7	309	21	AAW06065	Arabidopsis thaliana
32	34	60.7	309	21	AAW22955	Arabidopsis thaliana
33	34	60.7	368	22	AAW96392	Human papillomavir
34	34	60.7	368	22	AAW98400	Human papillomavir
35	34	60.7	368	23	ABP07669	Wild-type HPV6B E2
36	34	60.7	368	23	ABP07670	Mutant HPV6B E2 pr
37	34	60.7	389	21	AAW06064	Arabidopsis thaliana
38	34	60.7	389	21	AAW22954	Arabidopsis thaliana
39	34	60.7	389	23	ABP91402	Herbicidally activ
40	34	60.7	425	21	AAW06063	Arabidopsis thaliana
41	34	60.7	430	23	ABP39867	Staphylococcus epi
42	34	60.7	453	23	AAW22953	Arabidopsis thaliana
43	34	60.7	679	24	ABJ19484	UL15 DNA packaging
44	34	60.7	679	24	ABJ19529	UL15 DNA packaging
45	34	60.7	1587	18	AAW01816	Human papillomavir

ALIGNMENTS

RESULT 1

AAW50177

ID AAW50177 standard, peptide; 11 AA.

AC AAW50177;

XX

DT 16-JUL-1998 (first entry)

XX

DE Signal-transducing protein carboxy-terminal peptide.

XX

KW Inhibition; specific binding; signal-transducing protein;

KW cytoplasmic protein; proliferation; cancer cell; apoptosis;

KW virally infected cell.

XX

OS Synthetic.

XX

PN WO9805347-A1.

XX

PD 12-FEB-1998.

XX

PF 18-JUL-1997; 97WO-US12677.

XX

PR 22-JUL-1996; 96US-0681219.

XX

PA (UYCO) UNIV COLUMBIA NEW YORK.

XX

PI Sato T, Yanagisawa J;

XX

DR WPI; 1998-145347/13.

XX

PT Inhibition of signal transduction - by inhibiting binding between a signal-transducing protein and a cytoplasmic protein, for treating

PT e.g. cancer or viral infection
 XX
 PS Claim 19; Page 64; 108bp; English.
 CC A novel composition is capable of inhibiting specific binding
 CC between a signal-transducing protein (STP) having the
 CC carboxy-terminal sequence (Ser/Thr)-Xaa-(Val/Ile/Leu), where Xaa =
 CC any amino acid (e.g. the present peptide), and a cytoplasmic
 CC protein (CP) containing the sequence AAW50162 or AAW50163.
 CC The composition can be used to inhibit the proliferation of cancer
 CC or virally infected cells, or induce apoptosis in cancer or virally
 CC infected cells.
 CC
 SQ Sequence 11 AA;
 Query Match 100.0%; Score 56; DB 19; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.00044;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QNFRRTYVSFV 11
 DB 1 QNFRRTYVSFV 11
 RESULT 2
 ABP25867
 ID ABP25867 standard; Protein; 203 AA.
 AC ABP25867;
 DT 02-JUL-2002 (first entry)
 DE Streptococcus polypeptide SEQ ID NO 910.
 XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KM antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX Streptococcus pyogenes.
 OS
 XX
 PN WO200234771-A2.
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB04789.
 XX
 PR 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 PI Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tettelein H;
 XX
 DR MPI: 2002-352536/38.
 DR N-PSDB; ABN66498.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 PS Claim 1; Page 3244; 4525bp; English.
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 CC
 SQ Sequence 203 AA;
 Query Match 66.1%; Score 37; DB 23; Length 203;
 Best Local Similarity 60.0%; Pred. No. 39;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 QNFRRTYVSF 10
 DB 16 QNFRRTYVSF 25
 RESULT 3
 ABP81504
 ID ABP81504 standard; Protein; 850 AA.
 AC ABP81504;
 DT 04-MAR-2003 (first entry)
 DE Streptococcus pneumoniae polypeptide SEQ ID NO 421.
 XX Streptococcus pneumoniae; infection; otitis media; antibacterial;
 KW diagnosis; gene therapy.
 XX Streptococcus pneumoniae.
 OS
 XX
 PN WO200283855-A2.
 PD 24-OCT-2002.
 XX
 PF 12-APR-2002; 2002WO-US11524.
 XX
 PR 16-APR-2001; 2001US-283948P.
 PR 18-APR-2001; 2001US-284443P.
 XX
 PA (AMCY) AMERICAN CYANAMID CO.
 XX
 PI Zagursky RJ, Mast AW, Green BA, Chakravarti DN, Russell DP;
 PI Woocers JL;
 XX
 DR MPI: 2003-093010/08.
 DR N-PSDB; ABZ42352.
 XX
 PT New Streptococcus pneumoniae polynucleotides, useful for treating or
 PT preventing S. pneumoniae infections, or non-systemic diseases, e.g.
 PT otitis media, which are induced or exacerbated by S. pneumoniae -
 XX
 PS Claim 42; Page 696-699; 1091bp; English.
 CC The invention relates to isolated polynucleotides (ABZ72147-ABZ42352) of
 CC a Streptococcus pneumoniae genomic sequence, a fragment or degenerate
 CC variant of the polynucleotide or a nucleic acid sequence 95% identical to
 CC one of the polynucleotides. The S. pneumoniae polynucleotides and
 CC encoded polypeptides (ABP81299-ABP81674) are useful for treating or
 CC preventing S. pneumoniae infections or non-systemic diseases, e.g. otitis
 CC media, which are induced or exacerbated by S. pneumoniae. These are also
 CC useful for detecting S. pneumoniae in a biological sample or diagnosing
 CC S. pneumoniae infection in a subject. The polynucleotides have
 CC antibacterial activity and are useful in gene therapy.
 CC
 SQ Sequence 850 AA;
 Query Match 66.1%; Score 37; DB 24; Length 850;

Best Local Similarity 45.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 QNRTYIVSFV 11
: : : : :
Db 6 KTYRTYFISFI 16

RESULT 4
ABU02762
ID ABU02762 standard; Protein; 850 AA.

AC ABU02762;

DT 11-FEB-2003 (first entry)

DE 5. pneumoniae type 4 strain protein from coding region #2341.

XX Bacterial meningitis; pneumonia; sepsis; otitis media;

KW ear infection; antinflammatory; antibacterial; immunostimulant;

KM auditory; respiratory; gene therapy; vaccine.

OS Streptococcus pneumoniae type 4 strain.

PN W0200277021-A2.

PD 03-OCT-2002.

PF 27-MAR-2002; 2002WO-1B02163.

XX 27-MAR-2001; 2001GB-0007658.

XX (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

PI Maignani V, Tettelin H, Fraser C;

DR WPI, 2003-040579/03.

DR N-PSDB; ABX08053.

PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
useful as medicaments for treating or preventing a disease or infection
due to streptococcus bacteria, such as pneumonia, sepsis, otitis media
or ear infection

XX Claim 1; SEQ ID No 4682; 56pp; English.

CC The invention relates to a protein comprising or having at least 50%
identity to any of the 2469 amino acid sequences, identified in the
CC specification (available on a computer readable format), or its fragment,
CC expressed from 2469 of 2489 identified DNA coding regions from the
CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
CC AB56454. Also included are an antibody which binds one of the
CC proteins, treating a patient by administering the protein, DNA or
CC antibody (in a composition), a kit comprising first and second primers,
CC which are the nucleic acid cited above or fragments between nucleotides
CC 8-100 of a sequence not defined in the specification, for amplifying a
CC target sequence contained within a Streptococcus nucleic acid sequence,
CC where the first primer is substantially complementary to the target
CC sequence and the second primer is substantially complementary to the
CC complement of the target sequence, and where the parts of the primers
CC having substantial complementarity define the termini of the target
CC sequence to be amplified, assay comprising contacting a test compound
CC with the protein, and determining whether the test compound binds to the
CC protein and a Streptococcus pneumoniae bacterium, where one or more
CC genes encoding the proteins has been rendered inactive. The proteins,
CC nucleic acid molecules, antibody and compositions are useful as
CC medicaments for treating or preventing a disease or infection due to
CC streptococcus bacteria, particularly S. pneumoniae, such as pneumonia,
CC sepsis, otitis media or ear infection. They are also useful in developing
CC vaccines, diagnostic and antibiotics. The methods are useful for
CC identifying immunodominant proteins. The present sequence is one of
CC the 2469 proteins expressed by the identified coding regions from the

CC genomic sequence.

CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 850 AA;

Query Match 66.1%; Score 37; DB 24; Length 850;
Best Local Similarity 45.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 QNRTYIVSFV 11
: : : : :
Db 6 KTYRTYFISFI 16

RESULT 5
AAG42749
ID AAG42749 standard; Protein; 142 AA.

AC AAG42749;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 53350.

XX Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KM termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132484.

XX 06-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132487.

XX 11-MAY-1999; 99US-0132863.

XX 14-MAY-1999; 99US-0132856.

XX 14-MAY-1999; 99US-0134216.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 18-MAY-1999; 99US-0134370.

XX 18-MAY-1999; 99US-0134768.

XX 19-MAY-1999; 99US-0134941.

XX 20-MAY-1999; 99US-0135124.

XX 21-MAY-1999; 99US-0135353.

XX 24-MAY-1999; 99US-0135629.

XX 25-MAY-1999; 99US-0136021.

XX 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142855.
PR 06-JUL-1999; 99US-0142890.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0148668.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 64.3%; Score 36; DB 21; Length 142;
Best Local Similarity 70.0%; Pred. No. 41;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 QNFRYIVSF 10
|||
23 QNLSFIVSF 32
Db

RESULT 6
AAG55302
ID AAG55302 standard; Protein; 142 AA.
XX
AC AAG55302;
XX
DT 18-OCT-2000 (first entry)
XX
DB Arabidopsis thaliana protein fragment SEQ ID NO: 70887.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134841.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135323.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142290.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 64.3%; Score 36; DB 21; Length 142;
Best Local Similarity 70.0%; Pred. No. 41;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QNFRTYIVSF 10
|||:||||

Db 23 QNLSRTYIVSF 32
RESULT 7
ID AAM94340
ID AAM94340 standard; Protein; 154 AA.
XX
XX AAM94340;
AC
XX 21-NOV-2001 (first entry)
DT
XX
XX Human reproductive system related antigen SEQ ID NO: 2998.
DE
XX Human reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy.
XX
XX Homo sapiens.
OS
XX
XX WO200155320-A2.
PN
XX
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01339.
PF
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216847.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228927.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.

PR	08-SEP-2000	2000US-0232080.	PR	17-NOV-2000	2000US-0249300.	CC	The present invention provides the protein and coding sequences of a
PR	08-SEP-2000	2000US-0232081.	PR	01-DEC-2000	2000US-0250160.	CC	number of human reproductive system related antigens. These can be used
PR	12-SEP-2000	2000US-0231968.	PR	01-DEC-2000	2000US-0250391.	CC	in the prevention and treatment of reproductive system disorders,
PR	14-SEP-2000	2000US-0232397.	PR	05-DEC-2000	2000US-0251030.	CC	including cancer. The present sequence is a protein of the invention.
PR	14-SEP-2000	2000US-0232398.	PR	05-DEC-2000	2000US-0251988.	XX	
PR	14-SEP-2000	2000US-0232399.	PR	05-DEC-2000	2000US-0256719.	XX	
PR	14-SEP-2000	2000US-0232400.	PR	06-DEC-2000	2000US-0251479.	XX	
PR	14-SEP-2000	2000US-0232401.	PR	08-DEC-2000	2000US-0251856.	XX	
PR	14-SEP-2000	2000US-0233063.	PR	08-DEC-2000	2000US-0251868.	XX	
PR	14-SEP-2000	2000US-0233064.	PR	08-DEC-2000	2000US-0251869.	XX	
PR	14-SEP-2000	2000US-0233065.	PR	08-DEC-2000	2000US-0251989.	XX	
PR	21-SEP-2000	2000US-0234222.	PR	08-DEC-2000	2000US-0251990.	XX	
PR	21-SEP-2000	2000US-0234274.	PR	11-DEC-2000	2000US-0254097.	XX	
PR	25-SEP-2000	2000US-0234997.	PR	05-JAN-2001	2001US-0259678.	XX	
PR	25-SEP-2000	2000US-0234998.	PA	(HUMA-) HUMAN GENOME SCI INC.		XX	
PR	26-SEP-2000	2000US-0235484.	XX			XX	
PR	27-SEP-2000	2000US-0235834.	PI	Rosen CA, Baraesh SC, Ruben SM,		XX	
PR	27-SEP-2000	2000US-0235835.	XX	WPI: 2001-465570/50.		XX	
PR	29-SEP-2000	2000US-0236327.	DR	N-PSDB; AAL00310.		XX	
PR	29-SEP-2000	2000US-0236368.	XX			XX	
PR	29-SEP-2000	2000US-0236369.	PT	Isolated nucleic acid molecule encoding a reproductive system antigen		XX	
PR	29-SEP-2000	2000US-0236370.	PT	is used in preventing, treating or ameliorating a medical condition		XX	
PR	02-OCT-2000	2000US-0236802.	XX			XX	
PR	02-OCT-2000	2000US-0237037.	PS	Claim 11; SEQ ID NO 2998; 1297bp + Sequence Listing; English.		XX	
PR	02-OCT-2000	2000US-0237038.	CC			XX	
PR	02-OCT-2000	2000US-0237039.	CC	The present invention provides the protein and coding sequences of a		XX	
PR	13-OCT-2000	2000US-0239933.	CC	number of human reproductive system related antigens. These can be used		XX	
PR	13-OCT-2000	2000US-0239937.	CC	in the prevention and treatment of reproductive system disorders,		XX	
PR	20-OCT-2000	2000US-0240960.	CC	including cancer. The present sequence is a protein of the invention.		XX	
PR	20-OCT-2000	2000US-0241221.	XX			XX	
PR	20-OCT-2000	2000US-0241785.	SO	Sequence 154 AA;		XX	
PR	20-OCT-2000	2000US-0241786.				XX	
PR	20-OCT-2000	2000US-0241787.				XX	
PR	20-OCT-2000	2000US-0241808.				XX	
PR	20-OCT-2000	2000US-0241809.				XX	
PR	20-OCT-2000	2000US-0241826.				XX	
PR	01-NOV-2000	2000US-0244617.				XX	
PR	08-NOV-2000	2000US-0246474.				XX	
PR	08-NOV-2000	2000US-0246475.				XX	
PR	08-NOV-2000	2000US-0246476.				XX	
PR	08-NOV-2000	2000US-0246477.				XX	
PR	08-NOV-2000	2000US-0246478.				XX	
PR	08-NOV-2000	2000US-0246523.				XX	
PR	08-NOV-2000	2000US-0246524.				XX	
PR	08-NOV-2000	2000US-0246525.				XX	
PR	08-NOV-2000	2000US-0246526.				XX	
PR	08-NOV-2000	2000US-0246527.				XX	
PR	08-NOV-2000	2000US-0246528.				XX	
PR	08-NOV-2000	2000US-0246532.				XX	
PR	08-NOV-2000	2000US-0246609.				XX	
PR	08-NOV-2000	2000US-0246610.				XX	
PR	08-NOV-2000	2000US-0246611.				XX	
PR	08-NOV-2000	2000US-0246613.				XX	
PR	17-NOV-2000	2000US-0249207.				XX	
PR	17-						

PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142977.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.

PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 64.3%; Score 36; DB 21; Length 225;
Best Local Similarity 70.0%; Pred. No. 68;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QNFRITYVSP 10
DB 106 QNLSTRTVSP 115

RESULT 9
AAG55301
ID AAG55301 standard; Protein; 225 AA.
XX AAG55301;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 70886.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP103405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154339.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 64.3%; Score 36; DB 21; Length 225;

Best Local Similarity 70.0%; Pred. No. 68; Mismatches 1; Indels 2; Gaps 0;

Qy 1 QNRTYIVSF 10

Db 106 QNLSFTIVSF 115

RESULT 10
AA642747
ID AA642747 standard; Protein; 232 AA.

AC AA642747;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 53348.

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX EP1033405-A2.

PD 06-SEP-2000.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 64.3%; Score 36; DB 21; length 232;
Best Local Similarity 70.0%; Pred No. 70;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QNFRYIVSF 10
|||:||||
Db 113 QNLSTPIVSF 122

RESULT 11
AAG55300
ID AAG55300 standard; Protein; 232 AA.
XX AAG55300;
AC AAG55300;
XX AAG55300;
DF 18-OCT-2000 (first entry)
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 70885.
XX Arabidopsis thaliana protein fragment SEQ ID NO: 70885.
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX Arabidopsis thaliana.
PN EP1033405-A2.
XX EP1033405-A2.
PD 06-SEP-2000.
XX 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-2000; 2000EP-0301439.
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.

```

PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

```

```

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 64.3%; Score 36; DB 21; Length 232;
 Best Local Similarity 70.0%; Pred. No. 70;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 QNFRIVSVF 10
Db 113 QNLSRIVSVF 122

```

RESULT 12

```

ID ABB61754
ID ABB61754 standard; Protein; 366 AA.
AC ABB61754;
XX 26-MAR-2002 (first entry)
DT 26-MAR-2002 (first entry)
XX 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 12054.
XX Drosophila melanogaster polypeptide SEQ ID NO 12054.
KM Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
PD 27-SEP-2001.
XX 27-SEP-2001.
PF 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2001; 2001WO-US09231.
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE ) PE CORP NY.
PA Venter JC, Adams M, Li PWD, Myers EW;
PI WPI, 2001-658660/75.
DR N-PSDB; ABL05857.
XX N-PSDB; ABL05857.
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX interactions -
XX Disclousure; SEQ ID NO 12054; 21pp + Sequence Listing; English.
PS The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB857737-AB872072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 386 AA;

```

Query Match 64.3%; Score 36; DB 22; Length 386;
 Best Local Similarity 54.5%; Pred. No. 12e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 QNFRIVSVF 11
Db 308 QNFRIVSVF 318

```

```

RESULT 13
AB92124
ID ABB92124 standard; Protein; 510 AA.
XX
AC ABB92124;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 1335.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB ) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 1335; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (AB90790-AB99016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
SQ Sequence 510 AA;
XX
Query Match 64.3%; Score 36; DB 23; Length 510;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 3 FRTYIVSFV 11
|||::||
DB 3 FRAYIINFV 11
|||::||
RESULT 14
AB91234
ID ABB91234 standard; Protein; 510 AA.
XX
AC ABB91234;
XX
DT 27-FEB-2003 (first entry)
XX
DE Arabidopsis thaliana protein #62 modulated by PTGS.
XX
KM Posttranscriptional gene silencing; PTGS; plant; transformation.
XX
OS Arabidopsis thaliana.
XX
PN WO200281695-A2.

```

```

XX
PD 17-OCT-2002.
XX
PF 05-APR-2002; 2002WO-EP03806.
XX
PR 06-APR-2001; 2001US-282049P.
XX
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PA (FRIE-) FRIEDRICH MIESCHER INST.
XX
PI Zhu T, Glazov EA, Meine F, Wang X, Chang H;
XX
DR WPI; 2003-103337/09.
XX
DR N-PSDB; AB242078.
XX
PT Novel polynucleic acid segment useful for modulating gene expression
PT within a cell by posttranscriptional gene silencing, and for augmenting
PT a plant cell genome -
XX
PS Claim 53; Page 281-283; 438pp; English.
XX
CC The invention relates to a novel isolated polynucleic acid segment
CC modulated within a cell by posttranscriptional gene silencing (PTGS). The
CC invention specifically relates to a method to identify an expression
CC product that is modulated by PTGS. The polynucleotide is useful for
CC modulating the gene expression within a cell by PTGS, by introducing the
CC polynucleic acid into a cell and expressing the nucleic acid segment in
CC the cell to form a product. The polynucleic acid segment is also useful
CC for augmenting a cell genome, and for augmenting a plant genome, by
CC contacting a plant cell with the segment to produce a transformed plant
CC cell, and growing the transformed plant cell to produce a differentiated
CC transformed plant. The sequences shown in ABB81173 - ABB81298 represent
CC the product of a segment of A. thaliana cDNA modulated by PTGS.
XX
SQ Sequence 510 AA;
XX
Query Match 64.3%; Score 36; DB 24; Length 510;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 3 FRTYIVSFV 11
|||::||
DB 3 FRAYIINFV 11
|||::||
RESULT 15
ABB81581
ID ABB81581 standard; Protein; 1580 AA.
XX
AC ABB81581;
XX
DT 18-SEP-2002 (first entry)
XX
DE Human ABC-A-8-1 protein.
XX
KW Human; ABCA5; ABCA6; ABCA9; ABCA10; ATP-binding cassette transporter;
KW chromosome 17; chromosome 17q; chromosome 17q24; antiarteriosclerotic;
KW gene therapy; cholesterol; lipophilic molecule; inflammation;
KW prostaglandin; prostacyclin; arteriosclerosis; transport.
XX
OS Homo sapiens.
XX
PN WO200246458-A2.
XX
PD 13-JUN-2002.
XX
PF 07-DEC-2001; 2001WO-EP15401.
XX
PR 07-DEC-2000; 2000EP-0403440.
XX
PR 23-JAN-2001; 2001US-263231P.
XX
PA (AVET ) AVENTIS PHARMA SA.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

```

XX Deneffe P, Roelker-Montus M, Pradee C, Arnould-Reguigne I;
 PI Duverger N, Allikmets R, Dean M;
 XX MPI; 2002-557584/59.

XX A novel nucleic acid corresponding to ATP-binding cassette transporter
 PT gene and the encoded polypeptide, useful for preventing or treating a
 PT dysfunction in reverse transport of cholesterol

XX Disclosure; Fig 2; 216pp; English.

XX The present invention describes human ATP-binding cassette transporters
 CC (ABC). Specifically described are the human ABCA5, ABCA6, ABCA9 and
 CC ABCA10 genes (see ABN89594 to ABN89597) which encode the proteins given
 CC in ABN81574 to ABN81577). ABN89598 to ABN89715 represent ABCA5, ABCA6,
 CC ABCA9 and ABCA10 nucleotide fragments; and ABN89716 to ABN89806 represent
 CC primers for ABCA5, ABCA6, ABCA9 and ABCA10 genes which are used in the
 CC exemplification of the present invention. The ABC sequences have
 CC antiarteriosclerotic activities and can be used in gene therapy. ABC
 CC sequences can be used in the manufacture of a medicament intended for the
 CC prevention and/or treatment of a subject affected by a dysfunction in
 CC the reverse transport of cholesterol. The ABC proteins are involved in
 CC the reverse transport of cholesterol, in membrane transport of lipophilic
 CC molecules, in particular inflammation mediating substance such as
 CC prostaglandins and prostacyclins, or in any pathology whose candidate
 CC chromosomal region is situated on chromosome 17. They are also useful
 CC for the manufacture of a medicament intended for prevention of
 CC arteriosclerosis in various forms. The ABCA5, ABCA6, ABCA9 and ABCA10
 CC genes are located to chromosome 17, more specifically to the 17q24 locus.

XX Sequence 1580 AA:

Query Match 64.3%; Score 36; DB 23; Length 1580;
 Best Local Similarity 55.6%; Pred. No. 5.4e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 FRTYIVSPV 11

DB 1077 FMTYVISFI 1085

Search completed: November 26, 2003, 12:30:22
 Job time : 31.4819 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:22:05 ; Search time 5.30121 Seconds

(without alignments)
97.580 Million cell updates/sec

Title: US-09-230-111c-12
Perfect score: 56
Sequence: 1 QNFRYIVSFV 11

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	67.9	371	Y147_METUA	Q57611 methanococ
2	38	67.9	379	Y801_METUA	Q58211 methanococ
3	38	67.9	564	HCM1_YEAST	P25366 saccharomy
4	37	66.1	859	CHS1_SCHPO	P30597 schizosacch
5	36	64.3	147	Y720_SYNY3	P74650 bynchococyl
6	36	64.3	332	C1B3_CAVPO	Q94220 cavia porce
7	35	62.5	98	YJDK_ECOLI	P39275 escherichia
8	35	62.5	208	R1S8_BUCAP	Q8K422 buchnera ap
9	35	62.5	244	ATP6_BACSU	P37813 bacillus su
10	35	62.5	261	ATP6_SYNP6	P08444 synchococ
11	35	62.5	572	NUSM_LOCP1	Q36428 locusta mlg
12	35	62.5	653	NUOL_RICCP	Q92691 rickettsia
13	35	62.5	657	NUOL_RICCN	P35008 galdieria s
14	34	60.7	233	ATP1_GALSU	Q96860 oenothera h
15	34	60.7	247	ATP1_OENHO	Q84594 human papil
16	34	60.7	368	VE2_HPV6A	P03119 human papil
17	34	60.7	368	VE2_HPV6B	Q9B85 lotus japon
18	33	58.9	247	ATP1_LOTJA	P17344 zea mays (m
19	33	58.9	247	ATP1_MAIZE	Q96860 oenothera h
20	33	58.9	247	ATP1_MESVI	Q96860 oenothera h
21	33	58.9	247	ATP1_OEYSA	P12083 oryza sativ
22	33	58.9	247	ATP1_PRA	P06452 plasm sativ
23	33	58.9	247	ATP1_SEIOL	P06451 spinachia ol
24	33	58.9	247	ATP1_TOBAC	P06288 nicotiana t
25	33	58.9	247	ATP1_WHEAT	Q99200 triticum ae
26	33	58.9	248	ATP1_MARPO	P06589 marchantia
27	33	58.9	328	THM2_CAEEL	Q17330 caenorhabdi
28	33	58.9	477	SYC_CHITR	Q84877 chlamydia t
29	33	58.9	480	SYC_BORBU	Q84877 chlamydia t
30	33	58.9	712	Y1S3_HUMAN	O15060 homo sapien
31	33	58.9	886	Y1H3_YEAST	P40511 saccharomyc
32	33	58.9	933	CC54_YEAST	P30665 saccharomyc
33	32	57.1	154	MLA1_PIG	Q02781 sus scrofa

34	32	57.1	156	1	MLA1_RAT	P49218 rattus norv
35	32	57.1	195	1	BID_HUMAN	P55557 homo sapien
36	32	57.1	208	1	Y379_METUA	Q57611 methanococ
37	32	57.1	211	1	LOLA_COXBU	P39917 coxiella bu
38	32	57.1	238	1	ATP6_BACFI	P22476 bacillus fi
39	32	57.1	247	1	YG29_ARCFU	Q28644 archaeoglob
40	32	57.1	257	1	MLA1_BOVIN	Q02769 bos taurus
41	32	57.1	303	1	ATNB_PIG	P05027 sus scrofa
42	32	57.1	313	1	VP9_WTV	P12326 wound tumor
43	32	57.1	313	1	VP9_WTVNJ	P31611 wound tumor
44	32	57.1	333	1	C1B2_SHEEP	Q29422 ovine axes
45	32	57.1	350	1	MLA1_HUMAN	P48039 homo sapien

ALIGNMENTS

RESULT 1
ID Y147_METUA STANDARD; PRT; 371 AA.
AC Q57611;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical ATP-binding protein M0147.
GN M0147.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_Taxid=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Ueberback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Huret M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073 (1996).
RN [2]
RP SIMILARITY.
RX MEDLINE=97197912; PubMed=9045616;
RA Koonin E.V.;
RT "Evidence for a family of archaeal ATPases.";
RL Science 275:1489-1490 (1997).
CC -1- SIMILARITY: BELONGS TO THE ARCHAEL ATPASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U67472; AAB98139.1; -
CC PIR: D64318; D64318.
CC TIGR: M0147; -
CC InterPro: IPR003593; AAA ATPase.
CC InterPro: IPR002576; Archaeal ATPase.
CC Pfam: PF01637; Archaeal ATPase; 1.
CC ProDom: PD003608; Archaeal ATPase; 1.
CC SMART: SM00382; AAA; 1.
CC Hypothetical protein; ATP-binding; Complete proteome.
FT NP BIND 33 40 ATP (POTENTIAL).
SQ SEQUENCE 371 AA; 43859 MW; 088B8FC7D14F8B8 CRC64;

Query Match 67.9%; Score 38; DB 1; Length 371;
 Best Local Similarity 75.0%; Pred. No. 5.4;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NFRITYVS 9
 |||||:
 DB 61 NFRITYLIS 68

RESULT 2

Y801 METUA STANDARD; PRT; 379 AA.
 ID Y801 METUA
 AC 058211;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical ATP-binding protein W0801.
 GN W0801.
 OS Methanococcus jannaschii.
 OC Archaea: Euryarchaeota: Methanococci: Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus;
 OC NCBI_TaxId=2190;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 Sutton G.G., Blake O., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Getch C.L.,
 Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Gilek A.,
 Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
 Uitterlinden T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.,
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RT Science 273:1058-1073(1996).
 RL [2]

SIMILARITY.

RP MEDLINE=97197912; PubMed=9045616;
 RX Koonin E.V.;
 RA "Evidence for a family of archaeal ATPases";
 RT Science 275:1489-1490(1997).
 RL [1]

-1- SIMILARITY: BELONGS TO THE ARCHAEOAL ATPASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U67524; AAB98799.1; -
 CC DR PIR; A64400; A64400.
 DR TIGR; M0801; -
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR002576; Archaeal_ATPase.
 DR Pfam; PF01637; Archaeal_ATPase; 1.
 DR ProDom; PD003808; Archaeal_ATPase; 1.
 DR SMART; SM00382; AAA; 1.
 KW Hypothetical protein; ATP-binding; Complete proteome.
 FT NP_BIND 29 36 ATP (POTENTIAL)
 FT SEQUENCE 379 AA; 44716 MW; 1BAF2567E0C5D0B4 CRC64;

Query Match 67.9%; Score 38; DB 1; Length 379;
 Best Local Similarity 75.0%; Pred. No. 5.5;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NFRITYVS 9
 |||||:
 DB 57 NFRITYLIS 64

RESULT 3
 HCM1 YEAST STANDARD; PRT; 564 AA.
 ID HCM1 YEAST
 AC P25364;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE HCM1 protein.
 GN HCM1 OR YCR065W OR YCR65W OR YCR902.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OC NCBI_TaxId=4932;
 RN [1]

SEQUENCE FROM N.A.
 RP MEDLINE=93180825; PubMed=8441413;
 RX Zhu G., Muller E.G.D., Amacher S.L., Northrop J.L., Davis T.N.,
 RA "A dosage-dependent suppressor of a temperature-sensitive calmodulin
 RT mutant encodes a protein related to the fork head family of
 RT DNA-binding proteins";
 RL Mol. Cell. Biol. 13:1779-1787(1993).
 RN [2]

SEQUENCE FROM N.A.
 RP MEDLINE=92221691; PubMed=1561837;
 RA Bent P., Chanet R., Fabre F., Faye G., Fukuhara H., Sor F.,
 RT "Sequence of the supE1-RAD18 region on chromosome III of
 RT Saccharomyces cerevisiae";
 RL Yeast 8:147-153(1992).
 CC -1- FUNCTION: NOT KNOWN. DOSAGE DEPENDENT SUPPRESSOR OF A CALMODULIN
 CC MUTATION.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- SIMILARITY: Contains 1 fork-head domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; L08252; AAA34665.1; -
 CC DR EMBL; X59720; CAA42280.1; -
 DR PIR; S22262; S22262.
 DR HSP; Q63245; ZHRF.
 DR TRANSFAC; T01256; -
 DR SCD; S0000634; HCM1.
 DR GO; GO:0005634; C:nucleus; IPI.
 DR GO; GO:0003704; P:specific RNA polymerase II transcription fa. .; IDA.
 DR GO; GO:0007051; P:spindle assembly; IGI.
 DR GO; GO:0006367; P:transcription initiation from Pol II promoter; IDA.
 DR InterPro; IPR001766; TF_Fork_head.
 DR Pfam; PF00250; Fork_head; 1.
 DR ProDom; PD000425; TF_Fork_head; 1.
 DR SMART; SM00339; FH; 1.
 DR PROSITE; PS00657; FORK_HEAD_1; 1.
 DR PROSITE; PS00658; FORK_HEAD_2; 1.
 DR PROSITE; PS0039; FORK_HEAD_3; 1.
 KW DNA-binding; Nuclear protein.
 FT DNA_BIND 108 199 FORK-HEAD.
 FT CONFLICT 129 130 NV -> KL (IN REF. 2).
 FT CONFLICT 490 564 THEFQSPKSSAPDVTSTATSKFASGLFGVDYVYWR
 FT ATEKISDGNNTDSNQHPPYNNHNSDNGNEKN -> NFG
 FT NNFETKIVLCCTCPHRRNRFQICFKAVVRGCLFLEARN
 FT (IN REF. 2).
 SO SEQUENCE 564 AA; 63619 MW; 1A9D349F107B91BA CRC64;

Query Match 67.9%; Score 38; DB 1; Length 564;
 Best Local Similarity 70.0%; Pred. No. 8.2;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 QNFRITYVSP 10
 DB 352 QNFRKFTSF 361

RESULT 4
 CHS1_SCHPO STANDARD; PRT; 859 AA.

AC P30557;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chitin synthase 1 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl transferase 1)
 GN CHS1 OR SPAC13G6.12C OR SPAC24B11.01C.
 OS Schizosaccharomyces pombe (Pisleson yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 NCBI_TaxId=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Felkewell T., Fraser A., Gentile S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holtroyd S., Hornesby T., Howarth S., Huckle E.J., Hunt S., Jagsels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Ocell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E., Rutherford K., Rutter S., Saunders R., Squares S., Stevens K., Skelton J., Stammers M., Squares R., Squares K., Sharp S., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward G., Volckaert G., Aert R., Robben J., Grynoprez B., Weljens I., Vanstreels E., Rieger M., Schaefer M., Muller-Auer S., Gabel C., Fuchs M., Fitz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Leinhardt R., Pohl T.M., Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B., Goffeau A., Cadieu E., Drenth S., Gloux S., Lelaure V., Mottier S., Raebert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez L., Revelle J.L., Moreno S., Armstrong J., Foxburg S.L., Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J., Shpakovsky G.V., Ussery D., Barrett B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomycetes pombe."
 RL Nature 415:871-880(2002).
 RN [2]
 RP SEQUENCE OF 201-389 FROM N.A.
 RX MEDLINE=92115692; PubMed=1711323;
 RA Bowen A.R., Chen-Wu J.L., Momany M., Young R., Szaniszlo P.J., Robbing P.W.;
 RA "Classification of fungal chitin synthases."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:519-523(1992).
 CC -1- FUNCTION: PLAYS A MAJOR ROLE IN CELL WALL BIOGENESIS.
 CC -1- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + ((1,4)-(N-acetyl)-beta-D-glucosaminyl))((N)) = UDP + ((1,4)-(N-acetyl)-beta-D-glucosaminyl))((N+1)).
 CC -1- SUBCELLULAR LOCATION: Plasma membrane-bound.
 CC -1- SIMILARITY: BELONGS TO THE CHITIN SYNTHASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC EMBL; 254308; CAA91105.1; -

DR EMBL; 267757; CAA91766.1; -
 DR EMBL; M82957; AAA35297.1; -
 DR PIR; S62441; S62441.
 DR GeneDB; SPombe; SPAC13G6.12c; -
 DR InterPro; IPR004834; Chitin synth.
 DR InterPro; IPR001173; Glyco trans-2.
 DR Pfam; PF01644; Chitin synth. 1.
 DR ProDom; PD002998; Chitin synth. 1.
 KW Transferase; Glycosyltransferase; Transmembrane; Cell wall;
 KW Multigene family.
 FT TRANSMEM 544 564 POTENTIAL.
 FT TRANSMEM 615 635 POTENTIAL.
 FT TRANSMEM 662 682 POTENTIAL.
 FT TRANSMEM 793 813 POTENTIAL.
 FT TRANSMEM 853 853 POTENTIAL.
 FT TRANSMEM 853 853 POTENTIAL.
 FT CONFLICT 273 273 K -> R (IN REF. 2).
 SQ SEQUENCE 859 AA; 97989 MW; 63443713A7ADE421 CRC64;

Query Match
 Best Local Similarity 66.1%; Score 37; DB 1; Length 859;
 Best Local Similarity 75.0%; Pred. No. 19;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QNFRITYV 8
 DB 788 QNFRITYV 795

RESULT 5
 ID Y720 SYNY3 STANDARD; PRT; 147 AA.
 AC P74650;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein sl10720.
 GN SL10720.
 OS Synchocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
 NCBI_TaxId=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y., Miyajima N., Hikosawa M., Sugita A., Saito S., Kimura T., Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naro K., Okumura S., Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M., Tabata S.;
 RA "Sequence analysis of the genome of the unicellular cyanobacterium Synchocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions."
 RL DNA Res. 3:109-136(1996).
 CC -1- SIMILARITY: BELONGS TO THE RTX TOXIN ACYLTRANSFERASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC EMBL; D90917; BA18766.1; -
 DR PIR; S76854; S76854.
 DR InterPro; IPR003996; RtxC.
 DR Pfam; PF02794; HLYC; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 147 AA; 16960 MW; 4813A700816A4751 CRC64;

Query Match
 Best Local Similarity 64.3%; Score 36; DB 1; Length 147;
 Best Local Similarity 63.6%; Pred. No. 5.1;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 QNFRITYVSP 11

Db 29 QNPAKVPVSFI 39

```

RESULT 6
AC C1B3_CAVPO STANDARD; PRT; 332 AA.
ID C1B3_CAVPO
AC 090220;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD1b3 precursor (CD1-b3 antigen).
GN CD1B3.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hartley, and NIH 2; TISSUE=Thymus;
RX MEDLINE=20021645; PubMed=10553074;
RA Dascher C.C., Hironaka K., Naylor J.W., Brauer P.P., Brown K.A.,
RA Storey J.R., Behar S.M., Kawasaki E.S., Porcelli S.A., Brenner M.B.,
RA Leciair K.P.;
RA "Conservation of a CD1 multigene family in the guinea pig.";
RT J. Immunol. 163:5478-5488(1999).
CC -1- FUNCTION: NOT KNOWN.
CC -1- SUBUNIT: ASSOCIATES NON-COVALENTLY WITH BETA-2-MICROGLOBULIN (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AF145485; AAF12740.1; -.
DR HSSP; P11609; 1CD1.
DR InterPro; IPR007110; IG_1like.
DR InterPro; IPR003597; IG_1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00407; IgC1; 1.
DR PROSITE; PS50835; IG_1like; 1.
KW Glycoprotein; Signal; Transmembrane; Immunoglobulin domain;
KW Multigene family.
FT SIGNAL 1 17
FT CHAIN 18 332
FT DOMAIN 18 300
FT TRANSMEM 301 321
FT DOMAIN 322 332
FT DOMAIN 185 295
FT DISULFID 120 194
FT DISULFID 224 279
FT CARBOHYD 38 38
FT CARBOHYD 75 75
FT CARBOHYD 146 146
SQ SEQUENCE 332 AA; 37259 MW; AD2181E42AC6171A CRC64;

```

Query Match 64.3%; Score 36; DB 1; Length 332;
 Best Local Similarity 66.7%; Pred. No. 12;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 FRTYVSVF 11
 DB 88 FRTYVAVFI 96

RESULT 7

```

YUDK ECOLI
ID YUDK_ECOLI STANDARD; PRT; 98 AA.
AC P39275;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yj4k.
GN YUDK OR B4128 OR 25730 OR ECS5110.
OS Escherichia coli, and
OS Escherichia coli 0157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562; 83334;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RA "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119(1995).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RL O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U14003; AAA97028.1; -.
DR EMBL; AE000485; AAC70889.1; -.
DR EMBL; AE005646; AAG59327.1; -.
DR EMBL; AP002568; BAB38533.1; -.
DR PIR; C86108; C86108.
DR PIR; F91267; F91267.
DR PIR; S56357; S56357.
DR Ecogene; EG12468; yj4k.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 98 AA; 11468 MW; 8466741C825B468 CRC64;

```

Query Match 62.5%; Score 35; DB 1; Length 98;
 Best Local Similarity 75.0%; Pred. No. 5.3;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 FRTYVSVF 10
 DB 7 FRTYVSVF 14

```

RESULT 8
RISA_BUCAP          STANDARD;          PRT;          208 AA.
ID_RISA_BUCAP
AC_08KA22;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
GN Riboflavin synthase alpha chain (EC 2.5.1.9).
OS RIBB OR BUSG104.
OC Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxId=98794;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=22084549; PubMed=12089438;
RA Tamas I., Kleason L., Canhaeck B., Naeslund A.K., Eriksson A.-S.,
RA Wernegreen J.J., Sandstrom J.P., Moran N.A., Andersson S.G.E.;
RA "50 million years of genomic stasis in endosymbiotic bacteria.";
RL Science 296:2376-2379 (2002).
CC -1- FUNCTION: Riboflavin synthase is a bifunctional enzyme complex
CC catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-
CC ribityl-amino-2,4(1H,3H)-pyrimidinone and L-3,4-dihydroxy-2-
CC butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The alpha
CC subunit catalyzes the dimerization of 6,7-dimethyl-8-lumazine to
CC riboflavin and 5-amino-6-(1'-D)-ribityl-amino-2,4(1H,3H)-
CC pyrimidinone (By similarity).
CC -1- CATALYTIC ACTIVITY: 2 6,7-dimethyl-8-(1-D-ribityl)lumazine =
CC riboflavin + 4-(1-D-ribitylamino)-5-amino-2,6-dihydroxypyrimidine.
CC -1- PATHWAY: Final steps of riboflavin synthesis.
CC -1- SUBUNIT: Oligomer that consist of 3 alpha subunits and 60 beta
CC subunits (By similarity).
CC -1- SIMILARITY: TO BIOLUMINESCENCE ANTENNA PROTEINS LUXY (YFP) AND
CC LUXL (LUMP).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE014086; AAM67674.1; -
CC DR InterPro; IPR001783; Lum binding; 2.
CC DR Pfam; PF00677; Lum binding; 2.
CC DR ProDom; PD004110; Lum binding; 1.
CC DR TIGRPFAM; TIGR00187; F1be; 1.
CC DR PROSITE; PS00693; LUM_BINDING; 1.
CC KW Riboflavin biosynthesis; Transferase; Repeat; Complete proteome.
CC FT REPEAT 1 97 1.
CC FT SITE 81 85 2.
CC FT SITE 179 183 BINDS TO LUMAZINE (PROBABLE).
CC FT SITE 208 AA; 23364 MW; CE00E91AC732AF CRC64;
CC SQ SEQUENCE

```

```

Query Match          62.5%; Score 35; DB 1; Length 208;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 QNRTYVSF 10
DB 19 KNFHTYVNF 28

RESULT 9
ATP6_BACSU          STANDARD;          PRT;          244 AA.
ID_ATP6_BACSU
AC_P37813;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE ATP synthase A chain (EC 3.6.3.14) (Protein 6).
GN ATP6.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1423;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=168.
RX MEDLINE=95050246; PubMed=7961438;
RA Santana M., Ionescu M.S., Vertes A., Longin R., Kunst F., Danchin A.,
RA Glaeser P.;
RT "Bacillus subtilis FOE1 ATPase: DNA sequence of the atp operon and
RT characterization of atp mutants.";
RL J. Bacteriol. 176:6802-6811(1994).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Mooser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Bouter L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Comercon I.F., Cummings N.J., Daniel R.A.,
RA Dentzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Gallon N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollighly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert B., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karmata D., Kashara Y., Kleier-Blanchard M., Klein C.,
RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogawa A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Preecan E., Pujic P., Purnelle B., Rapoport G., Ray M., Reynolds S.,
RA Rieger M., Rivoita C., Rocha E., Roche B., Rose W., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scifone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
RA Tsuchi M., Tamakoshi A., Tanaka T., Tetsura P., Tognoni A.,
RA Toesto V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wamucut R., Wedler E., Wedler H., Melzinger T.,
RA Winieri P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis".
RL Nature 390:249-256(1997).
CC -1- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL, IT MAY PLAY A
CC DIRECT ROLE IN THE TRANSLLOCATION OF PROTONS ACROSS THE MEMBRANE.
CC -1- CATALYTIC ACTIVITY: ATP + H(2O) + H(+) (in) = ADP + phosphate +
CC H(+) (out).
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. CONTAINS 8
CC POTENTIAL TRANSMEMBRANE DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z28592; CAA82254.1; -
CC DR EMBL; Z29122; CAB15704.1; -
CC DR PIR; I40362; I40362.
CC HSP; P00855; ICI17.

```

DR Subtilisin; BG10815; atpb.
 DR Interpro: IPR000568; ATPynt_Asub.
 DR Pfam: PF00119; ATP-synt_A; 1.
 DR PRINTS; PR00123; ATPASEA.
 DR TIGRFAMs; TIGR01131; ATP_synt_6_or_A; 1.
 DR PROSITE; PS00449; ATPASE_A; 1.
 DR Hydrogen ion transport; CF(0); Transmembrane; Complete proteome.
 SQ SEQUENCE 244 AA; 27054 MW; E26172BA9F1AA248 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 244;
 Best Local Similarity 63.6%; Pred. No. 13;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ONFRTYVDFV 11
 DB 49 QNFMEYVDFV 59

RESULT 10
 ATP6_SYNPE

ID ATP6_SYNPE STANDARD; PRT; 261 AA.

AC P08444;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ATP synthase A chain (EC 3.6.3.14) (Protein 6).
 GN ATP OR ATP1.
 OS Synchococcus sp. (strain PCC 6301) (Anacystis nidulans).
 OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
 OX NCBI_TaxId=1139;
 RX MEDLINE=87311713; PubMed=3041005;
 RP SEQUENCE FROM N.A.

RA Cozens A.L., Walker J.E.
 RT "The organization and sequence of the genes for ATP synthase subunits in the cyanobacterium Synchococcus 6301. Support for an endosymbiotic origin of chloroplasts."

RT J. Mol. Biol. 194:359-383(1987).

RL J. Mol. Biol. 194:359-383(1987).

CC -1- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A DIRECT ROLE IN THE TRANSLLOCATION OF PROTONS ACROSS THE MEMBRANE.

CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate + H(+) (Out).

CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), Epsilon(1). CF(0) HAS THREE MAIN SUBUNITS: A, B AND C.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. CONTAINS 8 POTENTIAL TRANSMEMBRANE DOMAINS.

CC -1- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X05302; CAA28923.1; -

DR PIR; S10826; PWYCA6.

DR HSSP; P00855; 1C17.

DR Interpro: IPR000568; ATPynt_Asub.

DR Pfam: PF00119; ATP-synt_A; 1.
 DR PRINTS; PR00123; ATPASEA.
 DR TIGRFAMs; TIGR01131; ATP_synt_6_or_A; 1.
 DR PROSITE; PS00449; ATPASE_A; 1.
 DR Hydrogen ion transport; CF(0); Transmembrane.
 SQ SEQUENCE 261 AA; 28982 MW; 83BE87398D0888C CRC64;

Query Match 62.5%; Score 35; DB 1; Length 261;
 Best Local Similarity 45.5%; Pred. No. 14;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ONFRTYVDFV 11
 DB 77 QNFMEYVDFV 87

RESULT 11
 NUSM_LOCM1

ID NUSM_LOCM1 STANDARD; PRT; 572 AA.

AC Q36428;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
 GN NDS.

OS Locusta migratoria (Migratory locust).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
 OC Acridoidea; Acrididae; Oedipodinae; Locusta.

OX NCBI_TaxId=7004;
 RX MEDLINE=96139026; PubMed=8587138;
 RA Flook P.K., Rowell C.H.F., Gellissen G.;
 RT "The sequence, organization, and evolution of the Locusta migratoria mitochondrial genome."

RT J. Mol. Evol. 41:928-941(1995).
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X80245; CAA65338.1; -

DR Interpro: IPR003916; NADhub_oxrds.

DR Interpro: IPR001750; Oxidored_q1.

DR Pfam; PF00361; Oxidored_q1_N.

DR Pfam; PF00662; Oxidored_q1_N; 1.

DR PRINTS; PR01434; NADHDGNASE5.

KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.

SQ SEQUENCE 572 AA; 66045 MW; D846818F0BE8651 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 572;
 Best Local Similarity 50.0%; Pred. No. 31;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 NRTYVDFV 11
 DB 551 NRTYVDFV 560

RESULT 12
 NUOL_RICPR

ID NUOL_RICPR STANDARD; PRT; 653 AA.

AC Q9ZCG1;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NADH-quinone oxidoreductase chain L (EC 1.6.99.5) (NADH dehydrogenase I, chain L) (NDH-L, chain L).

GN NUOL OR RP792.

OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsia.

OX NCBI_TaxId=782;
 RX MEDLINE=782;
 RP SEQUENCE FROM N.A.

RA STRAIN=Madrid E;

RC

```

RX MEDLINE=9039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sticheritz-Ponten T., Almaraz U.C.M., Podowski R.M., Naslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria.";
RL Nature 396:133-140(1998).
CC -! FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
CC sulfur (Fe-S) centers, to quinones in the respiratory chain.
CC Couples the redox reaction to proton translocation (for every two
CC electrons transferred, four hydrogen ions are translocated across
CC the cytoplasmic membrane), and thus conserves the redox energy in
CC a proton gradient (by similarity).
CC -! CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -! SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -! SIMILARITY: TO POLYPEPTIDE 5 OF THE NADH-UBIQUINOL OXIDOREDUCTASE
CC OF CHLOROPLASTS OR MITOCHONDRIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL, AJ235273, CA15218.1; -.
CC DR PIR, B71640; B71640.
CC DR InterPro; IPR003916; NADhub_oxred5.
CC DR InterPro; IPR001750; Oxidored_q1.
CC DR InterPro; IPR001516; Oxidored_q1_N.
CC DR Pfam; PF00361; oxidored_q1; 1.
CC DR Pfam; PF00662; oxidored_q1_N; 1.
CC DR PRINTS; PR01434; NADHDHGAS55.
CC KW Oxidoreductase; NAD; Quinone; Transmembrane; Complete proteome.
CC FT TRANSMEM 8 28 POTENTIAL.
CC FT TRANSMEM 35 55 POTENTIAL.
CC FT TRANSMEM 89 109 POTENTIAL.
CC FT TRANSMEM 115 135 POTENTIAL.
CC FT TRANSMEM 138 158 POTENTIAL.
CC FT TRANSMEM 178 198 POTENTIAL.
CC FT TRANSMEM 214 234 POTENTIAL.
CC FT TRANSMEM 253 273 POTENTIAL.
CC FT TRANSMEM 285 305 POTENTIAL.
CC FT TRANSMEM 313 333 POTENTIAL.
CC FT TRANSMEM 336 356 POTENTIAL.
CC FT TRANSMEM 377 397 POTENTIAL.
CC FT TRANSMEM 418 438 POTENTIAL.
CC FT TRANSMEM 465 485 POTENTIAL.
CC FT TRANSMEM 511 531 POTENTIAL.
CC FT TRANSMEM 577 597 POTENTIAL.
CC FT TRANSMEM 629 649 POTENTIAL.
CC SQ SEQUENCE 653 AA; 73477 MW; 540436B913D097B4 CRC64;

Query Match
Query Local Similarity 62.5%; Score 35; DB 1; Length 653;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 NERTYIVSFV 11
DB 631 NVTLYIVSFI 640

RESULT 13
NUOL_RICCN STANDARD; PRT; 657 AA.
AC 092G57;
DT 28-FEB-2003 (rel. 41, Last sequence update)
DT 28-FEB-2003 (rel. 41, Last annotation update)
DB NADH-quinone oxidoreductase chain L (EC 1.6.99.5) (NADH dehydrogenase
DE 1, chain L) (NDH-L, chain L).
GN NUOL OR RC1226.

```

```

OS Rickettsia conorii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=781;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissensbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
CC -! FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
CC sulfur (Fe-S) centers, to quinones in the respiratory chain.
CC Couples the redox reaction to proton translocation (for every two
CC electrons transferred, four hydrogen ions are translocated across
CC the cytoplasmic membrane), and thus conserves the redox energy in
CC a proton gradient (by similarity).
CC -! CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -! SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -! SIMILARITY: TO POLYPEPTIDE 5 OF THE NADH-UBIQUINOL OXIDOREDUCTASE
CC OF CHLOROPLASTS OR MITOCHONDRIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL, AE008669; AL03764.1; ALT_INIT.
CC DR InterPro; IPR003916; NADhub_oxred5.
CC DR InterPro; IPR001750; Oxidored_q1.
CC DR InterPro; IPR001516; Oxidored_q1_N.
CC DR Pfam; PF00361; oxidored_q1; 1.
CC DR Pfam; PF00662; oxidored_q1_N; 1.
CC DR PRINTS; PR01434; NADHDHGAS55.
CC KW Oxidoreductase; NAD; Quinone; Transmembrane; Complete proteome.
CC FT TRANSMEM 5 25 POTENTIAL.
CC FT TRANSMEM 32 52 POTENTIAL.
CC FT TRANSMEM 89 109 POTENTIAL.
CC FT TRANSMEM 114 134 POTENTIAL.
CC FT TRANSMEM 138 158 POTENTIAL.
CC FT TRANSMEM 172 192 POTENTIAL.
CC FT TRANSMEM 214 234 POTENTIAL.
CC FT TRANSMEM 253 273 POTENTIAL.
CC FT TRANSMEM 285 305 POTENTIAL.
CC FT TRANSMEM 313 333 POTENTIAL.
CC FT TRANSMEM 339 359 POTENTIAL.
CC FT TRANSMEM 381 401 POTENTIAL.
CC FT TRANSMEM 416 436 POTENTIAL.
CC FT TRANSMEM 465 485 POTENTIAL.
CC FT TRANSMEM 511 531 POTENTIAL.
CC FT TRANSMEM 580 600 POTENTIAL.
CC FT TRANSMEM 632 652 POTENTIAL.
CC SQ SEQUENCE 657 AA; 73327 MW; 21326DC83CFE42DC CRC64;

Query Match
Query Local Similarity 62.5%; Score 35; DB 1; Length 657;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 NERTYIVSFV 11
DB 634 NVALYIVSFI 643

RESULT 14
ATPI_GALSU STANDARD; PRT; 233 AA.
AC P35008;
DT 01-FEB-1994 (rel. 28, Created)

```

DT 01-FEB-1994 (Rel. 28, last sequence update)
 DT 15-SEP-2003 (Rel. 42, last annotation update)
 DE ATP synthase A chain precursor (EC 3.6.3.14) (Subunit IV).
 GN ATP1.
 OS *Galdieria sulphuraria* (Red alga).
 OC Chloroplast.
 CC *Eukaryota*; *Rhodophyta*; *Bangiophyceae*; *Porphyridiales*; *Porphyridiaceae*;
 CC *Galdieria*.
 CC NCBI_Taxid=130081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=14-1-1 / Isolate 107.79/Goettingen;
 RX MEDLINE=94033298; PubMed=8219057;
 RA Koziczewa M., Zetsche K.;
 RT "organization of plastid-encoded ATPase gene and flanking regions
 including homologues of *infB* and *tsf* in the thermophilic red alga
Galdieria sulphuraria.";
 RT plant Mol. Biol. 23:67-76(1993).
 CC -1- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A
 DIRECT ROLE IN THE TRANSLLOCATION OF PROTONS ACROSS THE MEMBRANE.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
 H(+) (out).
 CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
 CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
 SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
 HAS THREE MAIN SUBUNITS: A, B AND C.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast
 CC thylakoid membrane.
 CC -1- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X67814; CAA48020.1; -
 DR HSSP: P00855; 1C17.
 DR InterPro: IPR000568; ATPsyn_Asub.
 DR Pfam: PF00119; ATP-synt_A; 1.
 DR PRINTS: PR00123; ATPASEA.
 DR TIGRFS: TIGR01131; ATP synt_6_or_A; 1.
 DR PROSITE: PS00449; ATPASE_A; 1.
 KW Hydrogen ion transport; CF(0); Chloroplast; Transmembrane; Signal.
 FT SIGNAL 1 43
 FT CHAIN 44 233 ATP SYNTHASE A CHAIN.
 FT TRANSMEM 82 102 POTENTIAL.
 FT TRANSMEM 121 141 POTENTIAL.
 FT TRANSMEM 190 210 POTENTIAL.
 FT TRANSMEM 211 231 POTENTIAL.
 SQ SEQUENCE 233 AA; 26066 MW; B02064124EB0A559 CRC64;

Query Match 60.7%; Score 34; DB 1; Length 233;
 Best Local Similarity 45.5%; Pred. No. 20;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 QNFRFYVSV 11
 DB 55 QNFRFYVSV 65

RESULT 15
 ID ATPI OENHO STANDARD; PRT; 247 AA.
 AC Q9MTM0;
 DT 15-SEP-2003 (Rel. 42, last sequence update)
 DT 15-SEP-2003 (Rel. 42, last annotation update)
 DE ATP synthase A chain precursor (EC 3.6.3.14) (Subunit IV).
 GN ATP1.
 OS *Oenothera hookeri* (Hooker's evening primrose).

OG Chloroplast.
 CC *Eukaryota*; *Viridiplantae*; *Streptophyta*; *Embryophyta*; *Tracheophyta*;
 CC *Spermatophyta*; *Magnoliophyta*; *Eudicotyledons*; *core eudicots*; *Rosidae*;
 CC *eurosid II*; *Myrtales*; *Onagraceae*; *Oenothera*.
 CC NCBI_Taxid=85636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Johansen;
 RX MEDLINE=20309318; PubMed=10852478;
 RA Hupfer H., Swiatek M., Hornung S., Herrmann R.G., Maier R.M.,
 RA Chiu W.-L., Sears B.;
 RT "Complete nucleotide sequence of the *Oenothera elata* plastid
 RT chromosome, representing plastome I of the five distinguishable
 RT *Eucnothera* plastomes.";
 RT Mol. Gen. Genet. 263:581-585(2000).
 CC -1- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A
 DIRECT ROLE IN THE TRANSLLOCATION OF PROTONS ACROSS THE MEMBRANE.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
 H(+) (out).
 CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
 CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
 SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
 HAS THREE MAIN SUBUNITS: A, B AND C.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast
 CC thylakoid membrane.
 CC -1- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AJ271079; CAB67156.1; -
 DR InterPro: IPR000568; ATPsyn_Asub.
 DR Pfam: PF00119; ATP-synt_A; 1.
 DR PRINTS: PR00123; ATPASEA.
 DR TIGRFS: TIGR01131; ATP synt_6_or_A; 1.
 DR PROSITE: PS00449; ATPASE_A; 1.
 KW Hydrogen ion transport; CF(0); Chloroplast; Transmembrane; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 247
 FT TRANSMEM 39 58 POTENTIAL.
 FT TRANSMEM 97 115 POTENTIAL.
 FT TRANSMEM 134 153 POTENTIAL.
 FT TRANSMEM 221 240 POTENTIAL.
 SQ SEQUENCE 247 AA; 27102 MW; 999709FE965BA1F9 CRC64;

Query Match 60.7%; Score 34; DB 1; Length 247;
 Best Local Similarity 54.5%; Pred. No. 21;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 QNFRFYVSV 11
 DB 69 QNFRFYVSV 79

Search completed: November 26, 2003, 12:31:16
 Job time : 6.30121 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:25:40 ; Search time 10.2048 Seconds

(without alignments)
103.662 Million cell updates/sec

Title: US-09-230-111C-12

Perfect score: 56

Sequence: 1 QNFRITYVSV 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	69.6	250	2 G81447	major antigenic pe
2	39	69.6	372	2 AH2214	hypothetical prote
3	38	67.9	371	2 D64318	hypothetical prote
4	38	67.9	379	2 A64400	hypothetical prote
5	38	67.9	491	2 C97069	cobryic acid synth
6	38	67.9	532	2 S22262	transcription fact
7	37	66.1	859	2 P95260	ABC transporter, P
8	37	66.1	859	2 S62441	chitin synthase (E
9	36	64.3	147	2 S76854	hypothetical prote
10	36	64.3	225	2 H96812	hypothetical prote
11	36	64.3	374	2 T25943	hypothetical prote
12	36	64.3	412	2 T46104	hypothetical prote
13	36	64.3	510	2 H84887	probable pectinest
14	36	64.3	1034	2 A95262	hypothetical prote
15	36	64.3	4981	2 T18489	hypothetical prote
16	35	62.5	98	2 F91267	hypothetical prote
17	35	62.5	98	2 C66108	hypothetical prote
18	35	62.5	98	2 S86357	hypothetical 11.5K
19	35	62.5	244	2 A10362	H+-transporting tw
20	35	62.5	261	1 PMYCA6	H+-transporting tw
21	35	62.5	346	2 A69993	hypothetical prote
22	35	62.5	572	2 T11478	NADH2 dehydrogenas
23	35	62.5	653	2 B71640	NADH2 dehydrogenas
24	35	62.5	659	2 AB1293	ABC transporter (p
25	35	62.5	660	2 B97853	ABC transporter (p
26	35	62.5	670	2 AH1664	ABC transporter (p
27	35	62.5	784	2 T45027	hypothetical prote
28	35	62.5	1808	2 T47792	hypothetical prote
29	35	62.5	2391	2 T18410	carbamoyl-phosphat

30	34	60.7	209	2 H70189	probable GDPdiacyl
31	34	60.7	211	2 S08522	probable sodium-de
32	34	60.7	233	2 S39515	H+-transporting tw
33	34	60.7	357	2 T11703	NADH2 dehydrogenas
34	34	60.7	368	1 B2WL6	E2 protein - human
35	34	60.7	376	2 AC2189	polysaccharide pol
36	34	60.7	389	2 B96635	hypothetical prote
37	34	60.7	527	2 T23951	hypothetical prote
38	34	60.7	565	2 F88362	protein R06A4.3 (1
39	34	60.7	613	2 B90294	hypothetical prote
40	34	60.7	646	2 B70396	histidine kinase s
41	34	60.7	674	2 T19495	hypothetical prote
42	33.5	59.8	279	2 G97143	probable membrane
43	33	58.9	87	2 H97778	hypothetical prote
44	33	58.9	195	2 A88360	protein W02B8.1 (1
45	33	58.9	216	2 T38520	hypothetical conse

ALIGNMENTS

RESULT 1
G81447
major antigenic peptide PEB3 Cj0289c [imported] - Campylobacter jejuni (strain NCTC 1116
C.Species: Campylobacter jejuni
C.Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C.Accession: G81447
R.Parkhill, J.; Wren, B.W.; Mungall, K.; Kestley, J.M.; Churcher, C.; Basham, D.; Chilling,
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A.Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy
A.Reference number: AB1250; MUID:20150912; PMID:10688204
A.Accession: G81447
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-250 <PAR-
A.Cross-references: GB:AL111168; NID:g6967505; PIDN:CAB72756.1; PID:g696776
A.Experimental source: Genotype O2, strain NCTC 11168
C.Genetics:
A.Gene: peb3; Cj0289c

Query Match 69.6%; Score 39; DB 2; Length 250;
Best Local Similarity 72.7%; Pred. No. 4.8;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 QNFRITYVSV 11
||| |||
||| |||
Db 158 QNFRNNIVAFV 168

RESULT 2
AH2214
hypothetical protein al13271 [imported] - Nostoc sp. (strain PCC 7120)
C.Species: Nostoc sp. PCC 7120
A.Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C.Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 03-Dec-2002
C.Accession: AH2214
R.Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Itiguchi,
N.; Kaneko, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, E
DNA Res. 8, 205-213, 2001
A.Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A.Reference number: AB1807; MUID:21595285; PMID:11759840
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-372 <KUR-
A.Cross-references: GB:BA000019; PIDN:BA074970.1; PID:g17132366; GSPDB:GN00179
A.Experimental source: strain PCC 7120
C.Genetics:
A.Gene: al13271
C.Superfamily: hypothetical protein yx8B

Query Match 69.6%; Score 39; DB 2; Length 372;

Best Local Similarity 77.8%; Pred. No. 7;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QNFRYYIVS 9
|||
Db 97 QNFRYYIVS 105

RESULT 3

D64318

hypothetical protein MJ0147 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: D64318

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,

Reich, C.T.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;

Reich, C.T.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;

Reich, C.T.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; PMID:96337999; PMID:8688087

A:Accession: D64318

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-371 <BUL>

A:Cross-references: GB:U67472; GB:L77117; NID:G2826253; PIDN:AAB98139.1; PID:G1590908; T

C:Genetics:

A:Map position: REV144813-143698

C:Superfamily: ATP-binding protein PAB1945

Query Match 67.9%; Score 38; DB 2; Length 371;

Best Local Similarity 75.0%; Pred. No. 11;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NFRYYIVS 9
|||
Db 61 NFRYYIVS 68

RESULT 4

A64400

hypothetical protein MJ0801 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: A64400

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,

Reich, C.T.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;

Reich, C.T.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;

Reich, C.T.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; PMID:96337999; PMID:8688087

A:Accession: A64400

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-379 <BUL>

A:Cross-references: GB:U67524; GB:L77117; NID:G2826321; PIDN:AAB98799.1; PID:G1499623; T

C:Genetics:

A:Map position: FOR725039-726178

C:Superfamily: ATP-binding protein PAB1945

Query Match 67.9%; Score 38; DB 2; Length 379;

Best Local Similarity 75.0%; Pred. No. 11;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NFRYYIVS 9
|||
Db 57 NFRYYIVS 64

RESULT 5

C97069

cobDytic acid synthase Cbp1 [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001

C:Accession: C97069

R:Molling, U.; Breton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J. Bacteriol. 183, 4823-4838, 2001

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C

A:Reference number: A96900; PMID:21359325; PMID:21359325

A:Accession: C97069

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-491 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK79342.1; PID:G15024310; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC1374

C:Superfamily: probable cobDytic acid synthase

Query Match 67.9%; Score 38; DB 2; Length 491;

Best Local Similarity 60.0%; Pred. No. 14;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NFRYYIVS 11
|||
Db 434 NFRYYIVS 443

RESULT 6

S22262

transcription factor HCM1 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YCR065w; protein YCR902

C:Species: Saccharomyces cerevisiae

C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 11-Jan-2002

C:Accession: S22262; S19480; A48153; S30779

R:Bent, P.; Chanet, R.; Fabre, F.; Faye, G.; Fukuhara, H.; Sor, F.

Yeast 8, 147-153, 1992

A:Title: Sequence of the supE1-PAD18 region on chromosome III of Saccharomyces cerevisiae

A:Reference number: S22260; PMID:92221691; PMID:1561837

A:Accession: S22262

A:Molecule type: DNA

A:Residues: 1-532 <BEN>

A:Cross-references: GB:X59720; EMBL:S93798; NID:G1907116; PIDN:CAA42280.1; PID:G1907205

R:Antoine, G.; Bent, P.; Chanet, R.; Faye, G.; Fukuhara, H.; Machieu, A.; So

submitted to the Protein Sequence Database, March 1992

A:Reference number: S19477

A:Accession: S19480

A:Molecule type: DNA

A:Residues: 1-532 <ANT>

A:Cross-references: EMBL:X59720; NID:G1907116; PIDN:CAA42280.1; PID:G1907205; MIPS:YCR06

R:Zhu, G.; Muller, E.G.; Amacher, S.L.; Northrop, J.L.; Davis, T.N.

Mol. Cell. Biol. 13, 1779-1787, 1993

A:Title: A dosage-dependent suppressor of a temperature-sensitive calmodulin mutant enc

A:Reference number: A48153; PMID:93180825; PMID:8441413

A:Accession: A48153

A:Molecule type: DNA

A:Residues: 1-489, 'TLETOISPRKSAADVLTATNSKFASSGLFGVDVSVWKRATEKISDGNNT', 'TDSNQKHPYHNP

A:Cross-references: EMBL:108252

A>Note: sequence extracted from NCBI backbone (NCBI:125990, NCBI:125991)

R:Zhu, G.; Muller, E.G.; Amacher, S.L.; Northrop, J.L.; Davis, T.N.

submitted to the EMBL Data Library, January 1993

A:Description: A dosage-dependent suppressor of a temperature-sensitive calmodulin mutan

A:Reference number: S30779

A:Accession: S30779

A:Molecule type: DNA

A:Residues: 1-128, 'NV', 131-489, 'TLETOISPRKSAADVLTATNSKFASSGLFGVDVSVWKRATEKISDGN', 'NTT

A:Cross-references: EMBL:108252; NID:G171653; PIDN:AAA34665.1; PID:G171654

C:Genetics:

A:Gene: SGD:HCN1

A:Cross-references: SGD:S0000661; MIPS:YCR065w

A:Map position: 3R

C:Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology

C:Keywords: DNA binding; nucleus

F.109-198/Domain: fork head DNA-binding domain homology <FHD>

Query Match 67.9%; Score 38; DB 2; Length 532;
Best Local Similarity 70.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 QNFRITYVSFV 10
|||||
352 QNFRKYFTSF 361

RESULT 7

AAC transporter, permease protein, probable SP2231 [Imported] - Streptococcus pneumoniae
P5260
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #ext_change 03-Aug-2001
C/Accession: P5260
R/Retcelin, H.; Nelson, K.E.; Paulsen, I.T.; Eissen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umeyam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,
neon, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A/Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A/Reference number: A95000; PMID:21357209; PMID:11463916
A/Accession: P5260
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-850 <KUR>
A/Cross-references: GB:M82957; PIDN:AAK76279.1; PID:G14973742; GSPDB:GN00164; TIGR:SP4
A/Experimental source: strain TIGR4
C/Genetics:
A/Genes: SP2231

Query Match 66.1%; Score 37; DB 2; Length 850;
Best Local Similarity 45.5%; Pred. No. 38;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 QNFRITYVSFV 11
:|||||:
Db 6 KTYRYFISFI 16

RESULT 8

CHitin synthase (EC 2.4.1.16) Chsi - fission yeast (Schizosaccharomyces pombe)
N/A/Alternate names: protein SPAC136.12c; protein SPAC24B11.01c
C/Species: Schizosaccharomyces pombe
C/Date: 16-May-1996 #sequence_revision 13-Mar-1997 #ext_change 20-Oct-2000
C/Accession: S62441; S62546; C45189; T37647; T38328
R/Odell, C.; Bowman, S.
Submitted to the EMBL Data Library, October 1995
A/Reference number: S62430
A/Accession: S62430
A/Residues: 1-859 <ODW>
A/Molecule type: DNA
A/Cross-references: EMBL:Z54308; NID:G1008985; PIDN:CAA91105.1; PID:G1008997
R/Odell, C.; Churcher, C.W.
Submitted to the EMBL Data Library, November 1995
A/Reference number: S62546
A/Accession: S62546
A/Molecule type: DNA
A/Residues: 1-199 <ODW>
A/Cross-references: EMBL:Z67757; NID:G1061288; PIDN:CAA91766.1; PID:G1061289
R/Bowen, A.R.; Chen-Wu, J.L.; Momany, M.; Young, R.; Szanialio, P.J.; Robbins, P.W.
Proc. Natl. Acad. Sci. U.S.A. 89, 519-523, 1992
A/Title: Classification of fungal chitin synthases.
A/Reference number: A38192; PMID:92115692; PMID:1731323
A/Accession: C45189
A/Molecule type: DNA
A/Residues: 201-272, 'R', 274-389 <BOW>
A/Cross-references: GB:M82957; NID:G173367; PIDN:AAA35297.1; PID:G173368
A/Note: sequence extracted from NCBI backbone (NCBIP:75847)
R/Odell, C.; Bowman, S.; Barrill, B.G.; Rajandream, M.A.; Walsh, S.V.; Wood, V.
Submitted to the EMBL Data Library, October 1995

A/Reference number: Z21734

A/Accession: T37647
A/Molecule type: DNA
A/Residues: 1-859 <OD2>
A/Cross-references: EMBL:Z54308; PIDN:CAA91105.1; GSPDB:GN00066; SPDB:SPAC1366.12c
A/Experimental source: strain 972h-; cosmid c13G6
R/Odell, C.; Churcher, C.W.; Barrill, B.G.; Rajandream, M.A.; Walsh, S.V.
Submitted to the EMBL Data Library, November 1995
A/Reference number: Z21786
A/Accession: T38328
A/Molecule type: DNA
A/Residues: 1-199 <OD3>
A/Cross-references: EMBL:Z67757; PIDN:CAA91766.1; GSPDB:GN00066; SPDB:SPAC24B11.01c
A/Experimental source: strain 972h-; cosmid c24B11
C/Genetics:
A/Genes: CHS1; SPAC24B11.01c; SPDB:SPAC1366.12c
A/Map position: 1L
A/Intons: 832/2
C/Superfamily: chitin synthase chsa

Query Match 66.1%; Score 37; DB 2; Length 859;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QNFRITYVSFV 8
:|||||:
Db 788 KNFRITYV 795

RESULT 9

hypothetical protein - Synecocystis sp. (strain PCC 6803)
S76854
C/Species: Synecocystis sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #ext_change 17-Feb-2003
C/Accession: S76854
R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasud
DNA Res. 3, 109-136, 1996
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocysti
s.
A/Reference number: S74322; PMID:97061201; PMID:8905231
A/Accession: S76854
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-147 <KAN>
A/Cross-references: EMBL:D90917; GB:AB001339; NID:G1653836; PIDN:BA18766.1; PID:d10194.
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C/Superfamily: hemolysin C

Query Match 64.3%; Score 36; DB 2; Length 147;
Best Local Similarity 63.6%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 QNFRITYVSFV 11
|||||
Db 29 QNFRKYFTSF 39

RESULT 10

hypothetical protein T30F21.21 [Imported] - Arabidopsis thaliana
H96812
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #ext_change 31-Mar-2001
C/Accession: H96812
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Comai, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luroe, J.S.; Maitl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

```

A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G., Frazer, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A6141; MUID:21016719; PMID:11130712
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-225 <STO>
A:Crosss-references: GB:AE005173; NID:g4836687; PIDN:AAD30590.1; GSPDB:GN00141
C:Genetics:
A:Gene: T30F21.21
A:Map position: 1

Query Match          64.3%; Score 36; DB 2; Length 225;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY      1  ONFRYYIVSF 10
          |||:|||||
Db      106  QNLSTFIYVSF 115

RESULT 11
T25943
hypoChemical protein ZC196.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25943
R:Murray, J.
submitted to the EMBL Data Library, April 1997
A:Description: The sequence of C. elegans cosmid ZC196.
A:Reference number: Z20115
A:Accession: T25943
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-374 <MUR>
A:Crosss-references: EMBL:U97007; PIDN:AAB52299.1; GSPDB:GN00023; CESP:ZC196.9
C:Genetics:
A:Experimental source: strain Bristol N2; clone ZC196
A:Gene: CESP:ZC196.9
A:Map position: 5
A:Introns: 7/1; 18/3; 54/3; 102/3; 143/1; 186/2; 247/3; 298/3; 326/3

Query Match          64.3%; Score 36; DB 2; Length 374;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY      2  NFRYYIVSF 10
          |||:|||||
Db      33  NFORYYIVSF 41

RESULT 12
T46104
hypoChemical protein T25B15.110 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 28-Jul-2000
C:Accession: T46104
R:Alcaraz, J.P.; Clabault, G.; Cottet, A.; Maché, R.; Mewes, H.W.; Lemcke, K.; Mayer, K
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23021
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-412 <ALC>
A:Crosss-references: EMBL:AL132972
A:Experimental source: cultivar Columbia; BAC clone T25B15
C:Genetics:
A:Map position: 3
A:Introns: 20/3; 155/3; 220/3; 317/1; 369/1; 384/3
A:Superfamily: Arabidopsis thaliana hypoChemical protein F24B22.230

```

Query Match Similarity 64.3% Score 36; DB 2; Length 412;
Best Local Similarity 60.0% Pred. No. 30;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 304 KMERYSASF 313

RESULT 13

QY 1 QNFRYYIVSF 10
: : ||| ||
: : ||| ||

probable pectinesterase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: H84887
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffett, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Unayam, L.; Tallon, L.;
Natus, D.; Nierman, W.C.; White, O.; Eisen, U.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A94420; MUID:20083487; PMID:110617157
A:Accession: H84887
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-510 <STO>
C:Cross-references: GB:AE002093; NID:92583131; PIDN:AA82640.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g45220
A:Map position: 2
C:Superfamily: pectinesterase

Query Match 64.3% Score 36; DB 2; Length 510;
Best Local Similarity 66.7% Pred. NO. 36;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 FRTYIVSFV 11
||| ||| : ||
||| ||| : ||

Db 3 FRAYIIINVF 11

RESULT 14

A95262
Probable formate dehydrogenase (EC 1.2.1.2) alpha chain fdoG [imported] - Sinorhizobium
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 17-May-2002
C:Accession: A95262
R: Barnett, M.T.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow
Proc. Natl. Acad. Sci. U.S.A. 98, 9893-9898, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: A95262
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1034 <KRR>
A:Cross-references: GB:AA006469; PIDN:AAK64659.1; PID:G14523056; GSPDB:GN00165
A:Experimental source: Strain 1021, megaplasmid pSymA
R: Gaibler, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler
Pela, D.; Chailu, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalmann, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure
hebaull, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: fdoG
A:Superfamily: formate dehydrogenase
C:Keywords: oxidoreductase

Query Match 64.3% Score 36; DB 2; Length 1034;

Best Local Similarity 60.0%; Pred. No. 71;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QNFRYIVSP 10
|||:|:|
Db 508 QNYRKFMVSP 517

RESULT 15

T18489
hypothetical protein C0820w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 09-Jun-2000
C:Accession: T18489
R:Lawson, D.; Bowman, S.; Barrell, B.
Submitted to the EMBL Data Library, August 1997
A:Reference number: Z18935
A:Accession: T18489
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-4981 <LAN>
A:Cross-references: EMBL:Z98551; NID:e1331903; PID:e1331910; PIDN:CAB11128.1
C:Genetics:
A:Map position: 3
A:Note: C0820w

Query Match 64.3%; Score 36; DB 2; Length 4981;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NFRTYIVSP 10
|||:|:|
Db 1536 NFQYIYTF 1544

Search completed: November 26, 2003, 12:36:01
Job time : 12.2048 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:25:10 / Search time 23.988 Seconds
(without alignments)
118.334 Million cell updates/sec

Title: US-09-230-111C-12
Perfect score: 56
Sequence: 1 QNFRYIVSFV 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_ricent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	71.4	364	2 Q8GFM0	Q8GFM0 staphylococ
2	40	71.4	489	12 Q92513	Q92513 beet virus
3	39	69.6	250	16 Q9PIK7	Q9PIK7 campylobact
4	39	69.6	372	16 Q8Y620	Q8Y620 anabaena sp
5	38	67.9	486	9 Q21944	Q21944 bacterioph
6	38	67.9	491	16 Q97J82	Q97J82 clostridium
7	38	67.9	576	16 Q8P2G3	Q8P2G3 streptococ
8	38	67.9	737	5 Q9N891	Q9N891 plasmid
9	37	66.1	312	5 Q9N891	Q9N891 plasmid
10	37	66.1	633	5 Q81DY3	Q81DY3 plasmid
11	37	66.1	716	12 Q91U09	Q91U09 inf. agent
12	37	66.1	850	16 Q97N40	Q97N40 streptococ
13	37	66.1	2190	12 Q915A5	Q915A5 pichinde ar
14	36	66.3	204	12 Q8QNN3	Q8QNN3 ectocarpus
15	36	66.3	225	10 Q9SYN6	Q9SYN6 arabidopsis
16	36	64.3	225	10 Q8L9P0	Q8L9P0 arabidopsis

17	36	64.3	272	10 Q8RWS3	Q8RWS3 arabidopsis
18	36	64.3	374	5 Q01625	Q01625 caenorhabdi
19	36	64.3	412	10 Q9FT48	Q9FT48 arabidopsis
20	36	64.3	413	5 Q9VL27	Q9VL27 drosophila
21	36	64.3	423	10 Q93XN8	Q93XN8 arabidopsis
22	36	64.3	427	5 Q8WYV3	Q8WYV3 drosophila
23	36	64.3	511	10 Q22149	Q22149 arabidopsis
24	36	64.3	687	11 Q8ROR4	Q8ROR4 mus musculu
25	36	64.3	744	11 Q8C0A9	Q8C0A9 mus musculu
26	36	64.3	970	11 Q8C114	Q8C114 mus musculu
27	36	64.3	1034	16 Q931E0	Q931E0 rhizobium m
28	36	64.3	1581	4 Q94911	Q94911 homo sapien
29	36	64.3	1620	11 Q8K442	Q8K442 mus musculu
30	36	64.3	1623	11 Q8K449	Q8K449 mus musculu
31	36	64.3	1624	11 Q8K441	Q8K441 mus musculu
32	36	64.3	4981	5 Q77372	Q77372 plasmid
33	35	62.5	239	8 Q9G012	Q9G012 proteobact
34	35	62.5	270	16 Q8RGJ9	Q8RGJ9 fusobacteri
35	35	62.5	293	16 Q8RBS1	Q8RBS1 thermomane
36	35	62.5	346	16 Q14505	Q14505 bacillus su
37	35	62.5	371	5 Q97311	Q97311 plasmid
38	35	62.5	413	10 Q9LUP2	Q9LUP2 arabidopsis
39	35	62.5	515	5 Q8SWR7	Q8SWR7 drosophila
40	35	62.5	548	5 Q9NEU2	Q9NEU2 caenorhabdi
41	35	62.5	659	16 Q8Y6E1	Q8Y6E1 listeria m
42	35	62.5	670	16 Q92AR2	Q92AR2 listeria m
43	35	62.5	682	5 Q9VJ95	Q9VJ95 drosophila
44	35	62.5	774	5 Q9VJ73	Q9VJ73 drosophila
45	35	62.5	1168	16 Q8EBJ1	Q8EBJ1 shewanelia

ALIGNMENTS

RESULT 1
Q8GFM0 PRELIMINARY; PRT; 364 AA.
ID Q8GFM0: 01-MAR-2003 (TRENBLrel. 23, Created)
AC Q8GFM0: 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE ORF18.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E-1;
RA Sugai M., Yamaguchi T., Hayashi T., Nakase K., Takami H.;
RT "Complete nucleotide sequence of Staphylococcus aureus E-1 EDINA
RT plasmid.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003089; BACS4510.1; -.
KW Plasmid.
SQ SEQUENCE 364 AA; 43904 MW; FB1157510D16FC18 CRC64;

Query Match Best Local Similarity 71.4%; Score 40; DB 2; Length 364;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 NFRYIVSFV 11
DB 137 NFRYIVSFV 146

RESULT 2
ID Q92513: PRELIMINARY; PRT; 489 AA.
AC Q92513: 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE Coat protein.
 OS Best virus O.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potomavirus.
 NCBI_TaxID=71972;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98378056; PubMed=9714254;
 RA Koenig R., Pfeij C., Beier C., Commandeur U.;
 RT "Genome properties of best virus O, a new furo-like virus from
 RT sugarbeet, determined from unpurified virus.";
 RL J. Gen. Virol. 79:2027-2036(1998).
 DR EMBL; AJ223597; CA11459.1; -
 DR InterPro; IPR001337; TMV coat.
 DR Pfam; PF00721; TMV coat; 1.
 SQ SEQUENCE 489 AA; 54507 MW; 87B93BFF186C052AF CRC64;

Query Match 71.4%; Score 40; DB 12; Length 489;
 Best Local Similarity 60.0%; Pred. No. 15;
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 QNFRITYVSF 10
 Db 351 QNFRITYLVNY 360

RESULT 3

Q9PIK7 PRELIMINARY; PRT; 250 AA.
 AC Q9PIK7;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Major antigenic peptide PEB3.
 GN PEB3 OR C0289C.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Campylobacteraceae; Campylobacter.
 OX NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
 RA Jagsels K., Kariyasev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Barrett B.G.;
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences.";
 RL Nature 403:665-668(2000).
 DR EMBL; AL139074; CAB72756.1; -
 KW Complete proteome.
 SQ SEQUENCE 250 AA; 27536 MW; B2A9AF0630A04065 CRC64;

Query Match 69.6%; Score 39; DB 16; Length 250;
 Best Local Similarity 72.7%; Pred. No. 12;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QNFRITYVSFV 11
 Db 158 QNFRNNIVAFV 168

RESULT 4

Q8YS20 PRELIMINARY; PRT; 372 AA.
 AC Q8YS20;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein A113271.
 GN A113271.
 OS Anabaena sp. (strain PCC 7120).

OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriiguchi M., Ishikawa K., Kimura T.,
 RA Kishida Y., Kohara S., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 DR EMBL; AP003592; BAB74970.1; -
 DR Hypothetical protein; Complete proteome.
 SQ SEQUENCE 372 AA; 43535 MW; 9D60D1D8311F0CF8 CRC64;

Query Match 69.6%; Score 39; DB 16; Length 372;
 Best Local Similarity 77.8%; Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QNFRITYVS 9
 Db 97 QNFRITYIS 105

RESULT 5

O21944 PRELIMINARY; PRT; 486 AA.
 ID O21944;
 AC O21944;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Glucosyl transferase II.
 GN GTRII.
 OS Bacteriophage SfII (Shigella flexneri bacteriophage II).
 OC Viruses.
 OX NCBI_TaxID=62284;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98086100; PubMed=9426131;
 RA Mavris M., Manning P.A., Morona R.;
 RT "Mechanism of bacteriophage SfII-mediated serotype conversion in
 RT Shigella flexneri.";
 RL Mol. Microbiol. 26:939-950(1997).
 DR EMBL; AF021347; AAC39273.1; -
 SQ SEQUENCE 486 AA; 55778 MW; C9D597EAF3CB20CB CRC64;

Query Match 67.9%; Score 38; DB 9; Length 486;
 Best Local Similarity 63.6%; Pred. No. 37;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 QNFRITYVSFV 11
 Db 163 QTFVYIIAFV 173

RESULT 6

Q97JB2 PRELIMINARY; PRT; 491 AA.
 ID Q97JB2;
 AC Q97JB2;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Cobylic acid synthase CblP.
 GN CAC1374.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1486;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;


```

RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omechenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.T.,
RA Tachson R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol 183:4823-4838(2001).
DR EMBL, AF007649; AAK79342.1; -.
DR InterPro; IPR000515; BPD_transp.
DR InterPro; IPR002586; CblA_P.
DR Pfam; PF01656; CblA; 1.
DR TIGRfam; TIGR00313; cobO; 1.
DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBER; 1.
DR Complete proteome.
SQ SEQUENCE 491 AA; 54530 MW; 08D0828BD4B28640 CRC64;

Query Match 67.9%; Score 38; DB 16; Length 491;
Best Local Similarity 60.0%; Pred. No. 38;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 NFRYIVSFV 11
DB 434 NFRYIVSFV 443

RESULT 7
Q9P2G3 PRELIMINARY; PRT; 576 AA.
ID Q9P2G3;
AC Q9P2G3;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical phage protein spym18_0369 (Hypothetical phage associated
DE protein spym3_0710).
GN SPYM18_0369 OR SPYM3_0710.
OS Streptococcus pyogenes (serotype M18), and
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_Taxid=186103, 198466;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=pyogenes; STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Mueser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=pyogenes; STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Mueser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10076-10083 (2002).
DR EMBL, AF009980; AAL97120.1; -.
DR InterPro; IPR005021; Phage_Termin.
DR Pfam; PF03354; Phage_Terminase; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 576 AA; 67816 MW; 4D64454743347AA3 CRC64;

```

```

Query Match 67.9%; Score 38; DB 16; Length 576;
Best Local Similarity 54.5%; Pred. No. 44;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 QNRYIVSFV 11
DB 65 ENFORIISFV 75

RESULT 8
O01772 PRELIMINARY; PRT; 737 AA.
ID O01772;
AC O01772;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein ZC581.3.
GN ZC581.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Felodertinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston B., Gaitung S., Le T.T.;
RT "The sequence of C. elegans cosmid ZC581.";
RL Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL, AF003134; AAB54141.4; -.
DR WormEpi; ZC581.3; CB31399.
KW Hypothetical protein.
SQ SEQUENCE 737 AA; 84361 MW; C8323AB05C173137 CRC64;

Query Match 67.9%; Score 38; DB 5; Length 737;
Best Local Similarity 70.0%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 NFRYIVSFV 11
DB 217 NFRDYFLSFV 226

RESULT 9
Q9N891 PRELIMINARY; PRT; 312 AA.
ID Q9N891;
AC Q9N891;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE Vir15 protein.
GN VIR15.
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid=5855;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=vivax; STRAIN=3D7 / Serotype 3D7;
RX del Portillo H.A., Lanzer M., Barrell B.G.;
RA del Portillo H.A., Lanzer M., Barrell B.G.;
RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL, AL360354; CAB96704.1; -.
SQ SEQUENCE 312 AA; 36805 MW; 2D2A08FBE4DEODCA CRC64;

Query Match 66.1%; Score 37; DB 5; Length 312;
Best Local Similarity 60.0%; Pred. No. 38;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 QNFRITYVSF 10
DB 132 ENFRNTVVSF 141

RESULT 10

081DY3 PRELIMINARY; PRT; 633 AA.
AC 081DY3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, last sequence update)
DE Hypothetical protein.
GN PF13_0192.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCB1_TaxID=63329;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.,
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844509; CAD52484.1; -
KW Hypothetical protein.
SQ SEQUENCE 633 AA; 75557 MW; CB3C43CBPFC4D865 CRC64;

Query Match 66.1%; Score 37; DB 5; Length 633;
Best Local Similarity 45.5%; Pred. No. 76;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 QNFRITYVSF 11
DB 572 KNFRITYVTYI 582

RESULT 11

091U09 PRELIMINARY; PRT; 716 AA.
AC 091U09;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)
DE RNA polymerase.
GN PA.
OS Influenza A virus (A/HongKong/97/98 (H5N1)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCB1_TaxID=88097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/Hong Kong/97/98;
RA Shaw M.W., Cooper L.A., Xu X., Thompson W.W., Krause S.L., Guan Y.,
RA Zhou N.N., Klimov A., Cox N.J., Webster R.G., Lim W., Shortridge K.F.,
RA Subbarao K.;
RT "Avian influenza A H5N1 and H9N2 viruses bearing a specific
RT constellation of nonglycoprotein genes caused illnesses in humans.";
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257202; AA49345.1; -
DR InterPro; IPR001009; RNA_pol_P2.
DR Pfam; PF00603; Flu_Pa_1.
SQ SEQUENCE 716 AA; 82699 MW; E6F31F77891EB48F CRC64;

Query Match 66.1%; Score 37; DB 12; Length 716;
Best Local Similarity 60.0%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 QNFRITYVSF 10
DB 227 ENFRITYVDGF 236

RESULT 12

097N40 PRELIMINARY; PRT; 850 AA.
AC 097N40;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
DE ABC transporter, permease protein, putative.
GN SP2231.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCB1_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtapple E., Khouri H., Wolf A.M., Ueberback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae";
RT Science 293:498-506(2001).
RL EMBL; AE007510; AAK76279.1; -
DR TIGR; SP2231; -
KW Complete proteome.
SQ SEQUENCE 850 AA; 97303 MW; 1ADED613F06B515 CRC64;

Query Match 66.1%; Score 37; DB 16; Length 850;
Best Local Similarity 45.5%; Pred. No. 18+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 QNFRITYVSF 11
DB 6 KTYRTYFISFI 16

RESULT 13

0915A5 PRELIMINARY; PRT; 2190 AA.
AC 0915A5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
DE L. polymerase.
GN L.
OS Pichinde arenavirus.
OC Viruses; ssRNA negative-strand viruses; Arenaviridae; Arenavirus.
OX NCB1_TaxID=11630;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AN3739;
RA Zhang L., Mariotti K.A., Harnish D.G., Aronson J.F.,
RT "Reassortant analysis of guinea pig virulence of Pichinde virus
RT variants.";
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF427517; AAL16099.1; -
DR InterPro; IPR007099; RNA_pol_NSvir.
DR PROSITE; PS50525; RDRP_Flu; I.
SQ SEQUENCE 2190 AA; 252091 MW; C16DF73E385B943 CRC64;

Query Match 66.1%; Score 37; DB 12; Length 2190;
Best Local Similarity 63.6%; Pred. No. 2.6e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 QNFRITYVSF 11
DB 628 QGFRYYIMAFV 638

```

RESULT 14
O80NN3 PRELIMINARY; PRT; 204 AA.
ID O80NN3
AC O80NN3
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Esv-1-28.
OS Ectocarpus siliculosus virus.
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phaeovirus.
OX NCBI_Taxid=37665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Esv-1;
RA Delarogue N., Bothe G., Pohl T., Knippers R., Mueller D.G., Boland W.;
RT "The complete nucleotide sequence of the Ectocarpus siliculosus virus
RT genome."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF204951; AK14454.1; -.
DR InterPro; IPR001214; SET.
DR SMART; SM00317; SET; 1.
SQ SEQUENCE 204 AA; 22852 MW; 03D757C3A8015EE1 CRC64;

Query Match 64.3%; Score 36; DB 12; Length 204;
Best Local Similarity 62.5%; Pred. No. 39;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 NFRITYVS 9
Db 61 NFRITYVS 68

RESULT 15
O9SYN6 PRELIMINARY; PRT; 225 AA.
ID O9SYN6
AC O9SYN6
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE T30F21.21 protein (F3F9.4).
GN T30F21.21
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Araujo R., Huitzer L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremenevskaja I., Kim C., Lenz C., Li J., Liu S.,
RA Luros S., Schwartz J., Shim P., Toriumi M., Vysotskaia V.S.,
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shin P., Altafi H., Bel Q., Chin C., Chlou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharaky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thavert A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F3F9 from chromosome
RT 1."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

```

```

RN [4]
RP SEQUENCE FROM N.A.
RC Cheuk R., Shin P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bel Q., Chin C., Chlou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharaky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thavert A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007260; AAD30590.1; -.
DR EMBL; AC013430; AAF71794.1; -.
DR InterPro; IPR006917; SOUL.
DR Pfam; PF04832; SOUL; 1.
SQ SEQUENCE 225 AA; 25338 MW; D0CA9754C381471C CRC64;

Query Match 64.3%; Score 36; DB 10; Length 225;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QNFRITYVSF 10
Db 106 QNFRITYVSF 115

Search completed: November 26, 2003, 12:34:32
Job time : 25.988 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 26, 2003, 12:26:20 : Search time 23.494 Seconds
(without alignments)
102.059 Million cell updates/sec

Title: US-09-230-111c-13

Perfect score: 64

Sequence: 1 SDSNNMMNELSEV 13

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 673684 seqs, 184443283 residues

Total number of hits satisfying chosen parameters: 673684

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	13	8 US-08-681-219-15	Sequence 15, Appl
2	64	100.0	13	12 US-10-092-138-13	Sequence 13, Appl
3	64	100.0	13	12 US-09-230-111c-13	Sequence 13, Appl
4	39	60.9	126	15 US-10-101-464A-816	Sequence 816, App
5	39	60.9	126	15 US-10-101-464A-905	Sequence 905, App
6	39	60.9	126	15 US-10-101-464A-955	Sequence 955, App
7	39	60.9	126	15 US-10-101-464A-977	Sequence 977, App
8	36	56.2	639	12 US-10-032-585-7531	Sequence 7531, App
9	35	54.7	200	11 US-09-759-130B-196	Sequence 196, App
10	35	54.7	337	11 US-09-759-130B-195	Sequence 195, App
11	35	54.7	348	12 US-09-882-227-218	Sequence 218, App
12	35	54.7	353	11 US-09-759-130B-193	Sequence 193, App
13	35	54.7	353	11 US-09-946-374-397	Sequence 397, App
14	35	54.7	353	12 US-10-015-387A-397	Sequence 397, App
15	35	54.7	353	12 US-10-006-130A-397	Sequence 397, App

16	35	54.7	353	12	US-10-006-1172A-397	Sequence 397, App
17	35	54.7	353	12	US-10-015-382A-397	Sequence 397, App
18	35	54.7	353	12	US-10-017-253A-397	Sequence 397, App
19	35	54.7	353	12	US-10-017-306A-397	Sequence 397, App
20	35	54.7	353	12	US-10-012-064A-397	Sequence 397, App
21	35	54.7	353	12	US-10-017-867A-397	Sequence 397, App
22	35	54.7	353	12	US-10-012-101B-397	Sequence 397, App
23	35	54.7	353	12	US-10-012-117A-397	Sequence 397, App
24	35	54.7	353	12	US-10-012-752A-397	Sequence 397, App
25	35	54.7	353	12	US-10-012-754A-397	Sequence 397, App
26	35	54.7	353	12	US-10-013-909A-397	Sequence 397, App
27	35	54.7	353	12	US-10-013-911A-397	Sequence 397, App
28	35	54.7	353	12	US-10-013-912A-397	Sequence 397, App
29	35	54.7	353	12	US-10-015-610A-397	Sequence 397, App
30	35	54.7	353	12	US-10-015-653A-397	Sequence 397, App
31	35	54.7	353	12	US-10-015-671A-397	Sequence 397, App
32	35	54.7	353	12	US-10-012-217A-397	Sequence 397, App
33	35	54.7	353	12	US-10-013-906A-397	Sequence 397, App
34	35	54.7	353	12	US-10-015-388A-397	Sequence 397, App
35	35	54.7	353	12	US-10-015-480A-397	Sequence 397, App
36	35	54.7	353	12	US-10-015-715A-397	Sequence 397, App
37	35	54.7	353	12	US-10-012-753A-397	Sequence 397, App
38	35	54.7	353	12	US-10-015-385A-397	Sequence 397, App
39	35	54.7	353	12	US-10-007-236A-397	Sequence 397, App
40	35	54.7	353	12	US-10-015-389A-397	Sequence 397, App
41	35	54.7	353	12	US-10-013-915A-397	Sequence 397, App
42	35	54.7	353	12	US-10-015-394A-397	Sequence 397, App
43	35	54.7	353	12	US-10-015-519A-397	Sequence 397, App
44	35	54.7	353	12	US-10-015-390A-397	Sequence 397, App
45	35	54.7	353	15	US-10-006-856A-397	Sequence 397, App

ALIGNMENTS

RESULT 1
US-08-681-219-15
Sequence 15, Application US/08681219
Publication No. US20020058607A1
GENERAL INFORMATION:
APPLICANT: Takaaki Sato and Junn Yanagisawa
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF
TITLE OF INVENTION: (PDZ/DHR) DOMAIN AND USES THEREOF
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/681,219
FILING DATE: 22-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/48962/JPM/JKM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-681-219-15

Query Match 100.0%; Score 64; DB 8; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDSNNMMNELSEV 13
|||||
Db 1 SDSNNMMNELSEV 13

RESULT 2
US-10-092-138-13
Sequence 13, Application US/10092138
Publication No. US20030170723A1
GENERAL INFORMATION:

APPLICANT: Sato, Taka-Aki
TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
FILE REFERENCE: 65823/JPW/PT
CURRENT APPLICATION NUMBER: US/10/092,138
CURRENT FILING DATE: 2002-09-06
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 13
LENGTH: 13
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial
US-10-092-138-13
Sequence:source:synthesized

Query Match 100.0%; Score 64; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDSNNMMNELSEV 13
|||||
Db 1 SDSNNMMNELSEV 13

RESULT 3

US-09-230-111C-13
Sequence 13, Application US/09230111C
Publication No. US20030203414A1
GENERAL INFORMATION:
APPLICANT: Yano, Taka-Aki
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN
TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGP (PDZ/DHR)
FILE REFERENCE: 48962-A-PCT-US
CURRENT APPLICATION NUMBER: US/09/230,111C
CURRENT FILING DATE: 1999-05-17
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 13
LENGTH: 13
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial
US-09-230-111C-13
Sequence:source:synthesized

Query Match 100.0%; Score 64; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDSNNMMNELSEV 13

Db 1 SDSNNMMNELSEV 13
|||||

RESULT 4
US-10-101-464A-916
Sequence 816, Application US/10101464A
Publication No. US20030046728A1
GENERAL INFORMATION:
APPLICANT: Strabala, Timothy
APPLICANT: Nieuwenhuizen, Nicolaas
APPLICANT: Higgins, Colleen M.
TITLE OF INVENTION: Compositions Isolated from Plant Cells
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
FILE REFERENCE: 11000.1020C2
CURRENT APPLICATION NUMBER: US/10/101,464A
CURRENT FILING DATE: 2002-03-18
PRIOR APPLICATION NUMBER: 09/704,302
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 09/228,986
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/162,866
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: PCT/US00/00724
PRIOR FILING DATE: 2000-01-11
NUMBER OF SEQ ID NOS: 989
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 816
LENGTH: 126
TYPE: PRT
ORGANISM: Eucalyptus grandis
US-10-101-464A-816
Sequence:source:synthesized

Query Match 60.9%; Score 39; DB 15; Length 126;
Best Local Similarity 72.7%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 DSNMMNELSE 12
|||||
Db 1 DSNMMNELSE 11

RESULT 5
US-10-101-464A-905

Sequence 905, Application US/10101464A
Publication No. US20030046728A1
GENERAL INFORMATION:
APPLICANT: Strabala, Timothy
APPLICANT: Nieuwenhuizen, Nicolaas
APPLICANT: Higgins, Colleen M.
TITLE OF INVENTION: Compositions Isolated from Plant Cells
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
FILE REFERENCE: 11000.1020C2
CURRENT APPLICATION NUMBER: US/10/101,464A
CURRENT FILING DATE: 2002-03-18
PRIOR APPLICATION NUMBER: 09/704,302
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 09/228,986
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/162,866
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: PCT/US00/00724
PRIOR FILING DATE: 2000-01-11
NUMBER OF SEQ ID NOS: 989
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 905
LENGTH: 412
TYPE: PRT
ORGANISM: Eucalyptus grandis
US-10-101-464A-905
Sequence:source:synthesized

Query Match 60.9%; Score 39; DB 15; Length 412;
Best Local Similarity 72.7%; Pred. No. 85;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 DSNMNNELSE 12
| | | | |
Db 1 DSNMNTSEASE 11

RESULT 6
US-10-101-464A-955

; Sequence 955, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Scribala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 955
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-955

Query Match 60.9%; Score 39; DB 15; Length 482;
Best Local Similarity 72.7%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 DSNMNNELSE 12
| | | | |
Db 71 DSNMNTSEASE 81

RESULT 7
US-10-101-464A-977

; Sequence 977, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Scribala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 977
; LENGTH: 997
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-977

Query Match 60.9%; Score 39; DB 15; Length 997;
Best Local Similarity 72.7%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 DSNMNNELSE 12
| | | | |
Db 586 DSNMNTSEASE 596

RESULT 8
US-10-032-585-7531

; Sequence 7531, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jians
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7531
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (639)..(639)
; OTHER INFORMATION: X=any amino acid
US-10-032-585-7531

Query Match 56.2%; Score 36; DB 12; Length 639;
Best Local Similarity 63.6%; Pred. No. 4.5e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 DSNMNNELSE 12
| | | | |
Db 353 DINVMNKLSE 363

RESULT 9
US-09-759-130B-196

; Sequence 196, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Frazer, Christopher C.
; APPLICANT: Sharp, John D.
; APPLICANT: Barnes, Thomas S.
; APPLICANT: Kirt, Susan J.
; APPLICANT: Mackay, Charles R.
; APPLICANT: Myers, Paul S.
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Wrighton, Nicolaas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE REFERENCE: MP100-5350NMIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24

```

; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 196
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-130B-196.

Query Match          54.7%; Score 35; DB 11; Length 200;
Best Local Similarity 53.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 SDSNNMMNLSRV 13
Db      118 SDVNLSHNLREV 130

RESULT 10
US-09-759-130B-195
; Sequence 195, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Frazer, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: MP100-5350NNIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
```

```

; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 195
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-130B-195

Query Match          54.7%; Score 35; DB 11; Length 337;
Best Local Similarity 53.8%; Pred. No. 3.3e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 SDSNNMMNLSRV 13
Db      118 SDVNLSHNLREV 130

RESULT 11
US-09-882-227-218
; Sequence 218, Application US/09882227
; Publication No. US20030158396A1
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oocmen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20030158396A1 Helicobacter Polypeptides in the
; FILE REFERENCE: 06132/047002
; CURRENT APPLICATION NUMBER: US/09/882,227
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/902,615
; PRIOR FILING DATE: 1997-07-29
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 218
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-882-227-218

Query Match          54.7%; Score 35; DB 12; Length 348;
Best Local Similarity 66.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 SDSNNMMNE 9
Db      63 SDLNMMNE 71

RESULT 12
US-09-759-130B-193
; Sequence 193, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Frazer, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
```



```
FILE REFERENCE: MP100-5350NMIM
CURRENT APPLICATION NUMBER: US/09/759,130B
PRIOR APPLICATION NUMBER: 2002-09-16
PRIOR FILING DATE: 2000-01-07
PRIOR APPLICATION NUMBER: US 09/479,249
PRIOR FILING DATE: 2000-01-07
PRIOR APPLICATION NUMBER: US 09/559,497
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 09/578,063
PRIOR FILING DATE: 2000-05-24
PRIOR APPLICATION NUMBER: US 09/333,159
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: US 09/596,194
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 09/342,364
PRIOR FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: US 09/608,452
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/393,996
PRIOR FILING DATE: 1999-09-10
PRIOR APPLICATION NUMBER: US 09/602,871
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 09/420,707
PRIOR FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 460
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 193
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
US-09-759-130B-193

Query Match      54.7% Score 35; DB 11; Length 353;
Best Local Similarity 53.8%; Pred.No. 3,4e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 SDSMMNNELSEV 13
DB      134 SDVNLSHNQLREV 146

RESULT 13
US-09-946-374-397
Sequence 397, Application US/09946374
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillen, Kenneth J.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tuma, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2830P1C1
CURRENT APPLICATION NUMBER: US/09/946,374
CURRENT FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: 60/098716
PRIOR FILING DATE: 1998-09-01
```

```
PRIOR APPLICATION NUMBER: 60/098723
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098749
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098750
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098803
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/098821
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/098843
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/099536
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099602
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099642
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099741
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099754
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099763
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099792
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099815
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100388
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100584
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100661
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100662
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100664
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100683
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100684
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100710
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100711
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100849
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100930
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101014
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101068
```

;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/101071
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/101279
;; PRIOR FILING DATE: 1998-09-22
;; PRIOR APPLICATION NUMBER: 60/101471
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101472
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101474
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101475
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101476
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101477
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101479
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101738
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101741
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101743
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101915
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101916
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/102207
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102240
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102307
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102330
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102331
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102484
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102487
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102570
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102571
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102684
;; PRIOR FILING DATE: 1998-10-01
;; PRIOR APPLICATION NUMBER: 60/102687
;; PRIOR FILING DATE: 1998-10-01
;; PRIOR APPLICATION NUMBER: 60/102965
;; PRIOR FILING DATE: 1998-10-02
;; PRIOR APPLICATION NUMBER: 60/103258
;; PRIOR FILING DATE: 1998-10-06
;; PRIOR APPLICATION NUMBER: 60/103314
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103315
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103328
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103395
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103396
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103401
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103449
;; PRIOR FILING DATE: 1998-10-06
;; PRIOR APPLICATION NUMBER: 60/103633
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103678
;; PRIOR FILING DATE: 1998-10-08

;; PRIOR APPLICATION NUMBER: 60/103679
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103711
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/104257
;; PRIOR FILING DATE: 1998-10-14
;; PRIOR APPLICATION NUMBER: 60/104987
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105000
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105002
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105104
;; PRIOR FILING DATE: 1998-10-21
;; PRIOR APPLICATION NUMBER: 60/105169
;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105266
;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105693
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105694
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105807

Query Match 54.7%; Score 35; DB 11; Length 353;
Best Local Similarity 53.8%; Pred. No. 3.4e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 SDSNNMMNELREV 13
DB 134 SDVNLSHNOJREV 146

RESULT 14
US-10-015-387A-397
; Sequence 397, Application US/10015387A
; Publication No. US20030135034A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Baton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gueney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC54
; CURRENT APPLICATION NUMBER: US/10/015,387A
; CURRENT FILING DATE: 2001-12-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 397
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-387A-397

Query Match 54.7%; Score 35; DB 12; Length 353;
Best Local Similarity 53.8%; Pred. No. 3.4e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 SDSNNMMNELREV 13
DB 134 SDVNLSHNOJREV 146

```
RESULT 15
US-10-006-130A-397
/ Sequence 397, Application US/10006130A
/ Publication No. US20030148375A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Botstein, David
/ APPLICANT: Desnovers, Luc
/ APPLICANT: Eaton, Dan I.
/ APPLICANT: Ferrara, Napoleone
/ APPLICANT: Fong, Sherman
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, Christopher J.
/ APPLICANT: Guirney, Austin L.
/ APPLICANT: Hillan, Kenneth J.
/ APPLICANT: Pan, James
/ APPLICANT: Paoni, Nicholas F.
/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
/ FILE OF INVENTION: Acids Encoding the Same
/ FILE REFERENCES: P2830P1C7
/ CURRENT APPLICATION NUMBER: US/10/006,130A
/ CURRENT FILING DATE: 2002-03-19
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 477
/ SEQ ID NO 397
/ LENGTH: 353
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-006-130A-397

Query Match          54.7%; Score 35; DB 12; Length 353;
Best Local Similarity 53.8%; Pred. No. 3.4e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 SDSNNMMNELSEY 13
DB      134 SDVNLISHNQLREV 146
```

Search completed: November 26, 2003, 12:38:39
Job time : 23.494 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using BW model

Run on: November 26, 2003, 12:21:29 ; Search time 36.0241 Seconds
(without alignments)
57.280 Million cell updates/sec

Title: US-09-230-111c-13
Perfect score: 64
Sequence: 1 SDSMMNNMSEV 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 156726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

```

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	13	19	AAW50176
2	42	65.6	1072	21	AAW50871
3	39	60.9	126	21	AAW5497
4	39	60.9	236	13	AAW24191
5	39	60.9	274	17	AAW14864
6	39	60.9	323	22	ABG24842
7	39	60.9	412	21	AAW25530
8	39	60.9	453	24	ABP79742
9	39	60.9	509	23	ABP65470

10	39	60.9	1005	22	ABG24846
11	37	57.8	118	21	AAW33343
12	37	57.8	314	21	AAW3198
13	37	57.8	380	21	AAW35197
14	37	57.8	408	22	ABW60374
15	37	57.8	420	21	AAW35196
16	37	57.8	1097	22	ABW57814
17	36	56.2	332	13	AAW21850
18	36	56.2	376	23	ABW51901
19	36	56.2	444	22	ABW08390
20	36	56.2	454	21	AAW24131
21	36	56.2	493	12	AAW13992
22	36	56.2	620	22	ABW15968
23	36	56.2	639	23	ABW13991
24	36	56.2	713	12	AAW33424
25	36	56.2	751	14	AAW33424
26	36	56.2	795	21	AAW03138
27	36	56.2	995	21	AAW03129
28	36	56.2	1027	14	AAW42203
29	36	56.2	1027	14	AAW43699
30	36	56.2	1040	24	ABW81250
31	36	56.2	1552	22	ABW71764
32	36	56.2	1984	22	ABW01338
33	35	54.7	93	23	ABW10600
34	35	54.7	96	23	ABW53163
35	35	54.7	125	21	ABW0441
36	35	54.7	125	23	ABW0541
37	35	54.7	143	21	AAW07838
38	35	54.7	146	20	AAW37267
39	35	54.7	159	22	AAW25942
40	35	54.7	207	23	ABW0685
41	35	54.7	252	21	AAW07837
42	35	54.7	253	21	AAW07836
43	35	54.7	272	20	AAW35675
44	35	54.7	313	21	AAW06775
45	35	54.7	313	21	AAW14540

ALIGNMENTS

RESULT 1
AAW50176 standard; peptide; 13 AA.

AAW50176;
16-JUL-1998 (first entry)

Signal-transducing protein carboxy-terminal peptide.

Inhibition; specific binding; signal-transducing protein;
cytoplasmic protein; proliferation; cancer cell; apoptosis;
virally infected cell.

Synthetic.

WO9805347-A1.
12-FEB-1998.

18-JUL-1997; 97WO-US12677.
22-JUL-1996; 96US-0681219.

(UYCO) UNIV COLUMBIA NEW YORK.

Sato T, Yanagisawa J;
WPI; 1998-145347/13.

Inhibition of signal transduction - by inhibiting binding between a
signal-transducing protein and a cytoplasmic protein, for treating

PT e.g. cancer or viral infection
XX
PS Claim 18; Page 64; 108bp; English.
XX
CC A novel composition is capable of inhibiting specific binding
CC between a signal-transducing protein (STP) having the
CC carboxy-terminal sequence (Ser/Thr)-Xaa-(Val/Ile/Leu), where Xaa =
CC any amino acid (e.g. the present peptide), and a cytoplasmic
CC protein (CP) containing the sequence AAW50162 or AAW50163.
CC The composition can be used to inhibit the proliferation of cancer
CC or virally infected cells, or induce apoptosis in cancer or virally
CC infected cells.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SDSNNMNNELSEV 13
Db 1 SDSNNMNNELSEV 13

RESULT 2
AAG70871
ID AAG70871 standard; Protein; 1072 AA.

AC AAG70871;
XX
DT 27-JUL-2001 (first entry)

DE C albicans apoptosis associated protein #51.

KW Yeast; fungus; apoptosis; infection; proliferative disease;
KW vaccine; autoimmune disease; ischaemia; neurodegeneration.

OS Candida albicans.

XX WO200102550-A2.

XX 11-JAN-2001.

PF 03-JUL-2000; 2000WO-BE00077.

XX 01-JUL-1999; 99EP-0870141.

XX (JANC) JANSSEN PHARM NV.

XX Contreras RH, De Backer MD, Luyten WHML, Malcorps IKL;
XX Nelissen BLM, Reekmans RJ;

XX WPI; 2001-367042/38.

XX N-PSDB; AAH29907.

XX Yeast and fungal nucleic acids encoding proteins involved in a pathway
PT leading to programmed cell death, useful for treating proliferative

PT disorders, yeast and fungal infections, or for preventing apoptosis in
PT certain diseases -

XX Claim 24; Fig 2; 218pp; English.

XX The present invention provides the protein and coding sequences of a
CC number of apoptosis associated proteins from the yeast *Saccharomyces*
CC *cerevisiae* and the fungus *Candida albicans*. These can be used to identify
CC treatments for fungal and yeast infections, for proliferative diseases
CC and for apoptosis related diseases such as autoimmune diseases, ischaemia
CC and neurodegeneration. The present sequence is one of the *C. albicans*
CC proteins of the invention.
XX
SQ Sequence 1072 AA;

Query Match 65.6%; Score 42; DB 22; Length 1072;

Best Local Similarity 53.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SDSNNMNNELSEV 13
Db 141 SDTDFNFELAEI 153

RESULT 3
AAB25497
ID AAB25497 standard; Protein; 126 AA.

AC AAB25497;

XX 27-NOV-2000 (first entry)

DE Eucalyptus grandis cell signalling involved protein SEQ ID NO:816.

KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
KW plant cell signalling; modulation; transgenic plant; pathogen; growth;

KW environmental change; development; cell proliferation; differentiation;
KW elongation; survival; disease resistance; nutrient metabolism.

XX Eucalyptus grandis.

XX WO200042171-A1.

XX 20-JUL-2000.

PF 11-JAN-2000; 2000WO-US00724.

XX 12-JAN-1999; 99US-0228986.

XX 01-NOV-1999; 99US-0162866.

XX (GENE) GENESIS RES & DEV CORP LTD.

XX Strabala TJ, Nieuwenhuizen NJ;

XX WPI; 2000-476052/41.

XX Isolated polynucleotide encoding a polypeptide involved in cell

PT signaling used for generating transgenic plants with modified responses
PT to external signals -

XX Claim 3; Page 389; 527pp; English.

XX AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
CC and protein sequences isolated from eucalyptus (*Eucalyptus grandis*) or
CC pine (*Pinus radiata* also known as Monterey pine). The protein sequences
CC are involved in cell signalling. The polynucleotide and protein
CC sequences can be used to modify the response of plant cells to external
CC signals e.g. environmental changes or pathogens during the growth and
CC development of a plant. They can be used to modify cell proliferation,
CC differentiation, elongation and survival, resistance to disease and
CC nutrient metabolism. Examples of modifications which can be produced are
CC altered fruit ripening and senescence of leaves and flowers e.g. to
CC delay senescence and prolong the life of cut flowers or enhance
CC senescence of reproductive organs to engineer sterile plants. Other
CC modifications can be used to delay senescence in selected cell types or
CC organs providing fruit and vegetables which have a longer shelf life
CC between harvest and consumption, or to decrease branching frequency in
CC forest tree species giving long stretches of valuable knot-free clear
CC wood which can be used in solid timber furniture and veneers.

XX SQ Sequence 126 AA;

Query Match 60.9%; Score 39; DB 21; Length 126;
Best Local Similarity 72.7%; Pred. No. 34;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 DSNMNNELSE 12
Db 1 DSNMNNELSE 11

```

RESULT 4
ID AAR24191 standard; Protein; 236 AA.
XX
AC AAR24191;
XX
DT 25-MAR-2003 (updated)
DT 24-NOV-1992 (first entry)
XX
DE Bovine RSV strain FS-1 P protein.
XX
KM Bovine respiratory syncytial virus; vaccine; diagnosis; antibodies;
KM P gene; phosphoprotein; BRSV.
XX
OS Bovine respiratory syncytial virus strain FS-1.
XX
PN WO9207940-A2.
XX
PD 14-MAY-1992.
XX
PF 04-NOV-1991; 91WO-US08177.
XX
PR 05-NOV-1990; 90US-0608937.
XX
PA (SAMA/) SAMAL S K.
XX
PI Samal SK;
XX
DR MPI; 1992-183675/22.
DR N-PSDB; AAQ25034.
XX
PT Bovine respiratory syncytial virus genes - used in the prodn. of
PT agents for use in detection and as vaccines for BRSV infection.
XX
PS Claim 33; Page 59; 74pp; English.
XX
CC This is the sequence of bovine respiratory syncytial virus (BRSV)
CC strain FS-1 phosphoprotein (P) protein. It can be used in the
CC detection of BRSV antibodies and in vaccines to prevent infection.
CC It can also be used for the production of BRSV protein antibodies.
CC See also AAR24185-R24190 and AAR25310.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 236 AA;

Query Match 60.9%; Score 39; DB 13; Length 236;
Best Local Similarity 46.2%; Pred. No. 72;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 DSNMMNMLESEV 13
   :|:|:|:|:|
   52 TSTNQINQLSEI 64

Db

RESULT 5
AAW14864
ID AAW14864 standard; protein; 274 AA.
XX
AC AAW14864;
XX
DT 16-MAY-1997 (first entry)
XX
DE Subtilisin DY variant A186N/S187D/F188Q/S190E.
XX
KM Subtilisin DY; variant; protease; cleaning composition; detergent.
XX
OS Synthetic.
XX
PN WO9628557-A2.
XX
PD 19-SEP-1996.

```

```

XX
PF 06-MAR-1996; 96WO-US03006.
XX
PR 09-MAR-1995; 95US-0401575.
XX
PA (PROC ) PROCTER & GAMBLE CO.
XX
PI Barnett BL, Brode PF, Rubingh DN;
XX
DR MPI; 1996-433829/43.
XX
PT Subtilisin DY variants having amino acid substitutions in loops 1-6
PT - are useful in cleaning compans, having decreased adsorption and
PT increased hydrolysis and giving improved cleaning performance.
XX
PS Examples 39-40; Page 131; 156pp; English.
XX
CC New subtilisin DY variants are provided in which one or more amino
CC acids within one or more of the six loop Regions have been replaced by
CC amino acids which are neutral or negatively charged and which are
CC equally or (preferably) more hydrophilic than the corresponding amino
CC acids in the wild-type sequence (see AAW03552).
CC Compared with wild-type Subtilisin DY, the variants have decreased
CC adsorption to, and provide increased hydrolysis of, insoluble protein
CC soils. They are useful as a component in a variety of cleaning
CC compositions. The present sequence is a specific example of a
CC variant enzyme suitable for use in a granular fabric cleaning
CC composition.
XX
SQ Sequence 274 AA;

Query Match 60.9%; Score 39; DB 17; Length 274;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 DSNMMNMLESEV 13
   |||:|:|:|
   180 DSNKNRNDQSEV 191

Db

RESULT 6
ABG24842
ID ABG24842 standard; Protein; 323 AA.
XX
AC ABG24842;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #24833.
XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYGE-) HYGEQ INC.
XX
PI Dymnac RT, Liu C, Tang YT;
XX
DR MPI; 2001-639362/73.
DR N-PSDB; AAS89029.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations

```

PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX
PS Claim 20; SEQ ID No 55201; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging or sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 323 AA;

Query Match 60.9%; Score 39; DB 22; Length 323;
Best Local Similarity 70.0%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SDSNNMMNML 10
: |||||:
DB 73 TQSNMMNMEV 82

RESULT 7
AAB25530
ID AAB25530 standard; Protein; 412 AA.
XX
XX AAB25530;
AC
XX
DT 27-NOV-2000 (first entry)
DE Eucalyptus grandis cell signalling involved protein SEQ ID NO:905.
XX
XX Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
KW environmental change; development; cell proliferation; differentiation;
KW elongation; survival; disease resistance; nutrient metabolism.
OS
XX Eucalyptus grandis.
OS
XX WO200042171-A1.
PN
XX 20-JUL-2000.
PD
XX 11-JAN-2000; 2000WO-US00724.
PF
XX 12-JAN-1999; 99US-0228986.
PR
XX 01-NOV-1999; 99US-0162866.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
PA
XX Strabala TJ, Nieuwenhuizen NJ;
PI
XX WPI; 2000-476052/41.
DR
XX Isolated polynucleotide encoding a polypeptide involved in cell
PT signalling used for generating transgenic plants with modified responses
PL to external signals -

XX
XX Claim 3; Page 466-467; 527bp; English.
PS
XX
XX AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
CC pine (Pinus radiata also known as Monterey pine). The protein sequences
CC are involved in cell signalling. The polynucleotide and protein
CC sequences can be used to modify the response of plant cells to external
CC signals e.g. environmental changes or pathogens during the growth and
CC development of a plant. They can be used to modify cell proliferation,
CC differentiation, elongation and survival, resistance to disease and
CC nutrient metabolism. Examples of modifications which can be produced are
CC altered fruit ripening and senescence of leaves and flowers e.g. to
CC delay senescence and prolong the life of cut flowers or enhance
CC senescence of reproductive organs to engineer sterile plants. Other
CC modifications can be used to delay senescence in selected cell types or
CC organs providing fruit and vegetables which have a longer shelf life
CC between harvest and consumption, or to decrease branching frequency in
CC forest tree species giving long stretches of valuable knot-free clear
CC wood which can be used in solid timber furniture and veneers.
XX
SQ Sequence 412 AA;

Query Match 60.9%; Score 39; DB 21; Length 412;
Best Local Similarity 72.7%; Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 DSNMMNMLSE 12
: |||||:
DB 1 DSNMTSNEASE 11

RESULT 8
ABP79742
ID ABP79742 standard; Protein; 453 AA.
XX
XX ABP79742;
AC
XX
DT 07-MAR-2003 (first entry)
DE N. gonorrhoeae amino acid sequence SEQ ID 6014.
XX
XX Antibacterial; infection; vaccine; gene therapy.
KW
XX
XX Neisseria gonorrhoeae.
OS
XX WO200279243-A2.
PN
XX 10-OCT-2002.
PD
XX 12-FEB-2002; 2002WO-1B02069.
PF
XX 12-FEB-2001; 2001GB-0003424.
PR
XX
XX (CHIR-) CHIRON SPA.
PA
XX Fontana MR, Pizsa M, Massignani V, Monaci E;
PI
XX WPI; 2003-058415/05.
DR
XX N-PSDB; ABZ40712.
DR
XX
PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection -
XX
XX Disclosure; Page 626; 815bp; English.
XX
XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABP79736-ABP81046 represent nucleic acid

CC molecules of the invention.
 XX
 SQ Sequence 453 AA;
 Query Match 60.9%; Score 39; DB 24; Length 453;
 Best Local Similarity 70.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SDSNNNNNEL 10
 : |||||:
 : |||||:
 Db 90 TSSNNNNNEV 99
 RESULT 9
 ABP65470
 ID ABP65470 standard; Protein; 509 AA.
 AC ABP65470;
 XX
 XX 19-NOV-2002 (first entry)
 DT
 XX Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:214.
 DE Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
 KM antidiarrhetic; antibacterial; inhibitor of Salmonella; detection;
 KM identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
 KM rotavirus; food composition; pharmaceutical composition.
 XX Bifidobacterium longum.
 OS
 XX EP1227152-A1.
 PN 31-JUL-2002.
 PD 30-JAN-2001; 2001EP-0102050.
 XX 30-JAN-2001; 2001EP-0102050.
 PR 30-JAN-2001; 2001EP-0102050.
 XX (NEST) SOC PROD NESTLE SA.
 PA
 XX WPI; 2002-668397/72.
 DR
 XX Novel polynucleotide comprising Bifidobacterium genome sequence useful
 PT as a probe or primer for detecting and/or identifying Bifidobacterium
 PT longum in a biological sample.
 PS Claim 3; SEQ ID 214; 80pp; English.
 XX
 XX The present invention describes a polynucleotide (I) comprising a
 CC sequence of a Bifidobacterium genome selected from the nucleotide
 CC sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at
 CC least 90% identity or which hybridises with the sequences given in
 CC ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding
 CC a fusion protein, comprising a sequence selected from 1097 sequences
 CC given in ABP65258 to ABP66354 ligated in frame to a polynucleotide
 CC encoding a heterologous polypeptide. (I) has antidiarrhetic and
 CC antibacterial activities, and can be used as an inhibitor of Salmonella.
 CC (I) (which is a probe) is useful for the detection and/or identification
 CC of Bifidobacterium longum in a biological sample. A carrier containing
 CC the lactic acid bacterium Bifidobacterium longum NCC2705 (CNCM I-2618)
 CC can be used for preventing and/or treating diarrhoea brought about by
 CC pathogenic bacteria and/or rotavirus. The carrier is a food composition
 CC selected from milk, yogurt, curd, cheese, fermented milks, milk based
 CC fermented products, ice-creams, fermented cereal based products, milk
 CC based powders, infant formula, pet food or a pharmaceutical composition
 CC selected from tablets, liquid bacterial suspensions, dried oral
 CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.
 CC (I) is useful in DNA arrays or chips to carry out analysis of the
 CC expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent
 CC Bifidobacterium related nucleotide sequences given in the Sequence
 CC listing from the present invention but not mentioned further within the
 CC specification.
 CC N.B. The sequence data for this patent is not represented in the printed

CC specification but is based on sequence information supplied by the
 CC European Patent Office.
 XX
 SQ Sequence 509 AA;
 Query Match 60.9%; Score 39; DB 23; Length 509;
 Best Local Similarity 70.0%; Pred. No. 1.8e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SDSNNNNNEL 10
 : |||||:
 : |||||:
 Db 138 TSSNNNNNEV 147
 RESULT 10
 ABG24846
 ID ABG24846 standard; Protein; 1005 AA.
 XX
 XX ABG24846;
 AC
 XX 18-FEB-2002 (first entry)
 DT
 XX Novel human diagnostic protein #24837.
 DE
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 OS
 XX WO200175067-A2.
 PN 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 XX 30-MAR-2001; 2001WO-US08631.
 PF 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 PI
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS89033.
 DR
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 PS Claim 20; SEQ ID No 55205; 103pp; English.
 XX
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1005 AA; 60.9%; Score 39; DB 22; Length 1005;
Query Match
Best Local Similarity 70.0%; Pred. No. 4e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 SDSNNNNNNEL 10
: |||||:
516 TQSNMNMNEV 525
Db
RESULT 11
AAG33343
ID AAG33343 standard; Protein; 118 AA.
AC AAG33343;
XX
XX 18-OCT-2000 (first entry)
XX DE Zea mays protein fragment SEQ ID NO: 40386.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
XX
XX Zea mays subsp. mays.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151203.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158433.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 57.8%; Score 37; DB 21; Length 118;
Best Local Similarity 46.2%; Pred. No. 71;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 SDSNNNNNELSEV 13
|||::|||::|:
Db 10 SDEWVSNMELMT 22

RESULT 12
AAG35198
ID AAG35198 standard; Protein; 314 AA.
XX
AC AAG35198;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 42963.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
OS Zea mays subsp. mays.
XX
XX PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140053.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149729.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 57.8%; Score 37; DB 21; Length 314;
Best Local Similarity 46.2%; Pred No. 2.2e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 SDSNNNNNEISEV 13
||::||::||:
Db 206 SDEWVSNNEMAEI 218

RESULT 13
AAG35197

ID AAG35197 standard; Protein; 380 AA.
XX AAG35197;
AC
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 42962.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence; corn.
OS Zea mays subsp. mays.
XX
XX EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2008EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139464.
PR 18-JUN-1999; 99US-0139465.
PR 21-JUN-1999; 99US-0139466.
PR 22-JUN-1999; 99US-0139467.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140355.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

```

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match          57.8%; Score 37; DB 21; Length 380;
Best Local Similarity 46.2%; Pred. No. 2.8e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 SDSNNNNNEISEV 13
DB 272 SDEWVSNMNMEMI 284

RESULT 14
ABB60374
XX ABB60374 standard; Protein; 408 AA.
AC ABB60374;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 7914.

```

```

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS WO200171042-A2.
XX
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW,
PI
XX MPI; 2001-656860/75.
DR N-PSDB; ABL04477.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 7914; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB57737-AB872072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 408 AA;

```

```

Query Match          57.8%; Score 37; DB 22; Length 408;
Best Local Similarity 58.3%; Pred. No. 3.1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 1 SDSNNNNNEISEV 12
DB 111 TOSNNNVNNESE 122

```

```

RESULT 15
AAG35196
ID AAG35196 standard; Protein; 420 AA.
XX
XX AAG35196;
AC
XX
DT 18-OCT-2000 (first entry)
XX
XX Zea mays protein fragment SEQ ID NO: 42961.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
XX Zea mays subsp. mays.
OS
XX EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-0301439.
XX

```

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132048.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 57.8%; Score 37; DB 21; Length 420;
Best Local Similarity 46.2%; Pred. No. 3.2e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPDSNMNMELSEV 13
DB 312 SDEMWSNMEMEI 324

Search completed: November 26, 2003, 12:30:24
Job time : 38.0241 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:22:05 ; Search time 6.26506 Seconds

(without alignments)
97.580 Million cell updates/sec

Title: US-09-230-111c-13

Perfect score: 64
Sequence: 1 SDSNNMMNELSEV 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length DB	ID	Description
1	40	62.5	1	BPPA_ECOL1
2	40	62.5	1	BPPA_ECOL1
3	39	60.9	1	FUMH_BACSU
4	39	60.9	1	FUMH_CAMJ2
5	39	60.9	1	FUMH_ECOL1
6	38	59.4	1	GNP1_YEAST
7	38	59.4	1	PEP3_YEAST
8	36	56.2	1	RELX_HORSE
9	36	56.2	1	THI1_SULSO
10	36	56.2	1	ASPA_HELPJ
11	36	56.2	1	ASPA_HELPJ
12	35	54.7	1	SECO_BORBU
13	35	54.7	1	VNS1_HRSVA
14	35	54.7	1	VNS1_HRSVA
15	35	54.7	1	P801_SCMA
16	35	54.7	1	ISCA_YEAST
17	35	54.7	1	RNHL_SCHPO
18	35	54.7	1	RMAR_CANCA
19	35	54.7	1	FUMH_HAEIN
20	35	54.7	1	Y045_WYCGE
21	35	54.7	1	Y045_WYCPN
22	35	54.7	1	K2C8_MOUSE
23	35	54.7	1	PPOD_LYCES
24	35	54.7	1	PPOB_LYCES
25	35	54.7	1	PPOC_LYCES
26	35	54.7	1	PPOC_LYCES
27	35	54.7	1	TRFE_BOVIN
28	35	54.7	1	YMA3_YEAST
29	35	54.7	1	PMH8_ARATH
30	35	54.7	1	CWFB_SCHPO
31	35	54.7	1	YGF0_YEAST
32	34	53.1	1	E417_ADE02
33	34	53.1	1	RELX_CANPA

ALIGNMENTS

34	34	53.1	209	1	SOMA_ESOLU	P34744 esox lucinus
35	34	53.1	310	1	YEG9_SCHPO	Q13771 schizosacch
36	34	53.1	339	1	SEH1_SCHPO	Q10099 schizosacch
37	34	53.1	615	1	CPO_DROME	Q01617 drosophila
38	34	53.1	619	1	DNAR_HELPJ	Q92mw4 helicobacte
39	34	53.1	619	1	DNAR_HELPJ	P55994 helicobacte
40	34	53.1	623	1	HAT0_CLOBO	P46085 clostridium
41	34	53.1	643	1	SGT1_ARATH	Q91ems arabidopsis
42	34	53.1	682	1	AC2L_MOUSE	Q99nd1 mus musculu
43	34	53.1	802	1	NAB3_YEAST	P38996 saccharomyc
44	34	53.1	899	1	PLC1_SCHPO	P40977 schizosacch
45	34	53.1	1848	1	CBPA_CLOCL	P38058 clostridium

RESULT 1

BPPA_ECOL1
ID BPPA_ECOL1 STANDARD, PRT, 193 AA.
AC P58997;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Major structural subunit of bundle-forming pilus precursor (Bundle-forming pilin) (Bundlin).
GN BPPA.
OS Escherichia coli O111:H-.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
OX NCB1_Taxid=168927;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O111:H- / B171;
RX MEDLINE=93211298; PubMed=8096320;
RA Sobel I., Puente J.L., Murray W.J., Vuopio-Varkila J., Schoolnik G.K.;
RT "Cloning and characterization of the bundle-forming pilin gene of enteropathogenic Escherichia coli and its distribution in Salmonella serotypes."
RT Mol. Microbiol. 7:563-575 (1993).
RL [2]
RP SEQUENCE OF 14-41.
RC STRAIN=O111:H- / B171;
RX MEDLINE=92054556; PubMed=1683004;
RA Giron J.A., Ho A.S.Y., Schoolnik G.K.;
RT "An inducible bundle-forming pilus of enteropathogenic Escherichia coli."
RT Science 254:710-713 (1991).
RL
CC -1- FUNCTION: Major repeating bundle-forming pilus (BFP) subunit. Is required for EPEC localized adherence.
CC -1- SUBUNIT: 10 to 100 laterally aligned filaments or bundle-forming pili coalesce into rope-like bundles. These form linkages between microcolonies that are attached to epithelial cells.
CC -1- SIMILARITY: BELONGS TO THE N-ME-PHE PILI FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.1eb-stb.ch/announce/> or send an email to license@1eb-stb.ch).
CC
CC EMBL: U27184; AAC4040.1;
CC InterPro: IPR001120; Prok_N_methyln.
DR Pfam: PF05307; Bundlin, 1.
DR PROSITE: PS00409; PROKAR_NTER_METHYL, 1.
KW Fibriin, Methylation, Plasmid.
FT PROPEP 1 13
FT CHAIN 14 193
FT MAJOR STRUCTURAL SUBUNIT OF BUNDLE-FORMING PILUS.

FT MOD RES 14 14 METHYLATION (POTENTIAL).
 FT DISULFID 129 179 POTENTIAL.
 FT CONFLICT 41 14 L -> M (IN REF. 2).
 FT CONFLICT 41 41 S -> I (IN REF. 2).
 SQ SEQUENCE 193 AA, 20300 MW, 2FCEC03B415FF026 CRC64;

Query Match 62.5%; Score 40; DB 1; Length 193;
 Best Local Similarity 61.5%; Pred. No. 3.6;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SDSNNMNNELSEV 13
 |||| : | : |||
 Db 41 SDSNKSQNAISEV 53

RESULT 2
 BPPA_ECO27 STANDARD; PRT; 193 AA.

AC P33553;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Major structural subunit of bundle-forming pilus precursor (bundle-forming pilin) (Bundlin).
 GN BPPA.
 OS Escherichia coli O127:H6.
 OC Plasmid pMAR2.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_Taxid=168807;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O127:H6 / E2348/69;
 RX MEDLINE=93133122; PubMed=1352446;
 RA Domeneberg M.S., Giron J.A., Nataro J.P., Kaper J.B.;
 RT "A plasmid-encoded type IV fimbrial gene of enteropathogenic
 RL Escherichia coli associated with localized adherence.";
 RM Mol. Microbiol. 6:3427-3437(1992).
 CC -1- FUNCTION: Major repeating bundle-forming pilus (BFP) subunit. Is
 CC required for BPEC localized adherence.
 CC -1- SUBUNIT: 10 to 100 laterally aligned filaments or bundle-forming
 CC pili coalesce into rope-like bundles. These form linkages between
 CC the bacteria within the enteropathogenic E.coli (BPEC)
 CC microcolonies that are attached to epithelial cells (by
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO THE N-ME-PHE PILI FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z12295; CAA78167.1; -
 CC PIR: S70966; S70966.
 DR PDB: 1OT2; 12-JUL-99.
 DR InterPro: IPR001120; Prok_N_methyln.
 DR Pfam: PF05307; Bundlin; 1.
 DR PROSITE: PS00409; PROKAR_NTER_METHYL; 1.
 KM Fimbria; Methylaton; Plasmid; 3D-structure.
 FT PROPEP 1 13 BY SIMILARITY.
 FT CHAIN 14 193 MAJOR STRUCTURAL SUBUNIT OF BUNDLE-
 FT FORMING PILUS.
 FT MOD RES 14 14 METHYLATION (POTENTIAL).
 FT DISULFID 129 179 POTENTIAL.
 SQ SEQUENCE 193 AA, 20270 MW, 2FCEC02E101AF426 CRC64;

Query Match 62.5%; Score 40; DB 1; Length 193;
 Best Local Similarity 61.5%; Pred. No. 3.6;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SDSNNMNNELSEV 13
 |||| : | : |||
 Db 41 SDSNKSQNAISEV 53

RESULT 3

FUMH_BACSU STANDARD; PRT; 462 AA.

AC P07343; O32194; (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Fumarate hydratase, class-II (EC 4.2.1.2) (fumarase).
 GN CING.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=85215459; PubMed=3923430;
 RA Miles J.S., Guest J.R.;
 RT "Complete nucleotide sequence of the fumarase gene (cing) of Bacillus
 RL subtilis 168.";
 RM Nucleic Acids Res. 13:131-140(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98304083; PubMed=9639930;
 RA Wipat A., Brignell C.S., Guy J.B., Rose M., Emerson P.T.,
 RA Harwood C.R.;
 RT "The yvra-yvga (293 degrees - 289 degrees) region of the Bacillus
 RT subtilis chromosome containing genes involved in metal ion uptake and
 RL a putative sigma factor.";
 RM Microbiology 144:1593-1600(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunet F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brunsch C.V., Caldwell I.F., Capuano N.J., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Eutlian K.D., Errington J., Fader C., Ferrari B., Fougere D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golligly E.J., Grand A.,
 RA Guseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kashara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porrolik S., Prescott A.M.,
 RA Priescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi Y., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi B., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpeira P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambuit R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumslein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RM Nature 390:249-256(1997).
 RL [4]
 RP SEQUENCE OF 1-50 FROM N.A.
 RC STRAIN=168;

```

RX MEDLINE=86056997; PubMed=2998952;
RA "Peavere I.M., Miles J.S., Moir A.;"
RT "The nucleotide sequence of a spore germination gene (gerA) of
RL Bacillus subtilis 168."
RN Gene 38:95-102(1985).
RP
RC
RX STRAIN=168;
MEDLINE=90036677; PubMed=2509423;
RA "Prie V.A., Peavere I.M., Moir A.;"
RT "Role of sigma H in expression of the fumarase gene (citG) in
RL vegetative cells of Bacillus subtilis 168."
RN Bacteriol. 171:5933-5939(1989).
RP
RN
RX SEQUENCE OF 224-462 FROM N.A.
MEDLINE=9801515; PubMed=9353911;
RA Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.;"
RT "Sequencing of regions downstream of addA (98 degrees) and citG (289
RL degrees) in Bacillus subtilis."
RU Microbiology 143:3305-3308(1997).
CC
CC -1- CATALYTIC ACTIVITY: (S)-malate + fumarate + H(2)O.
CC -1- SUBUNIT: Homotetramer.
CC -1- SIMILARITY: TO OTHER THERMOSTABLE CLASS II FUMARASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; X01701; CAA25849.1; -
DR EMBL; AJ223978; CAA11749.1; -
DR EMBL; Z99120; CAB15294.1; -
DR EMBL; M11918; AAA22464.1; -
DR EMBL; Z93941; CAB07973.1; -
DR PIR; A23033; UFBSC8.
DR HSSP; P05042; 1FUR.
DR Subtilisin; BG10384; citG.
DR InterPro; IPR005677; Fum_hydII.
DR InterPro; IPR000362; Fumarate_lyase.
DR Pfam; PF00206; lyase_1; 1.
DR PRINTS; PR00149; FUMARATE_LYASE.
DR TIGRFAMs; TIGR00979; func_II; 1.
DR PROSITE; PS00163; FUMARATE_LYASES; 1.
KW lyase; Tricarboxylic acid cycle; Complete proteome.
FT ACT SITE 186
FT BINDING 322 322 SUBSTRATE CARBOXYL (POTENTIAL).
FT CONFLICT 66 66 A -> V (IN REF. 1).
SQ SEQUENCE 462 AA; 50531 MW; 8D797D18FEA82C06 CRC64;

Query Match 60.9%; Score 39; DB 1; Length 462;
Best Local Similarity 70.0%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SDSNNANMNL 10
DB 99 TQSNMNMNEV 108

RESULT 4
FUNG_CAMJ2 STANDARD; PRT; 463 AA.
ID FUNG_CAMJ2
AC 069234;
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 28-SEP-2003 (Rel. 41; Last annotation update)
DE Fumarate hydratase class II (EC 4.2.1.2) (Fumarase)
GN FUNG OR CJ1364C.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;

```

```

OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=NCCT 11168;
RA Griffiths P.L., Connerton I.F.;"
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
RN
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=NCCT 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Felwell T., Holtroyd S.,
RA Jagers K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrett B.G.;"
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RL reveals hypervariable sequences."
CC Nature 403:665-668(2000).
CC
CC -1- CATALYTIC ACTIVITY: (S)-malate + fumarate + H(2)O.
CC -1- PATHWAY: Tricarboxylic acid cycle.
CC -1- SUBUNIT: Homotetramer (by similarity).
CC -1- SIMILARITY: TO OTHER THERMOSTABLE CLASS II FUMARASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; Y16882; CAA76499.1; -
DR EMBL; AL139078; CAB73791.1; -
DR PIR; A81281; A81281.
DR HSSP; P05042; 1FUR.
DR InterPro; IPR005677; Fum_hydII.
DR InterPro; IPR000362; Fumarate_lyase.
DR Pfam; PF00206; lyase_1; 1.
DR PRINTS; PR00149; FUMARATE_LYASE.
DR TIGRFAMs; TIGR00979; func_II; 1.
DR PROSITE; PS00163; FUMARATE_LYASES; 1.
KW lyase; Tricarboxylic acid cycle; Complete proteome.
FT ACT SITE 187
FT BINDING 323 323 SUBSTRATE CARBOXYL (POTENTIAL).
SQ SEQUENCE 463 AA; 50710 MW; E281B37B5791FFC0 CRC64;

Query Match 60.9%; Score 39; DB 1; Length 463;
Best Local Similarity 70.0%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SDSNNANMNL 10
DB 99 TQSNMNMNEV 108

RESULT 5
FUNG_ECOLI STANDARD; PRT; 467 AA.
ID FUNG_ECOLI
AC P05042; P76891;
DT 13-AUG-1987 (Rel. 05; Created)
DT 13-AUG-1987 (Rel. 05; Last sequence update)
DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE Fumarate hydratase class II (EC 4.2.1.2) (Fumarase)
GN FUNG OR B1611.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=87099873; PubMed=3541901;

```

RA Woods S.A., Miles J.S., Roberts R.E., Guest J.R.;
 RT "Structural and functional relationships between fumarase and
 RT aspartase. Nucleotide sequences of the fumarase (fumC) and aspartase
 RT (aspA) genes of *Escherichia coli* K12.";
 RL Biochem. J. 237:547-557(1986).
 RN [12]
 RP SEQUENCE FROM N.A.
 RP STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN [13]
 RP SEQUENCE FROM N.A.
 RP STRAIN=K12;
 RX MEDLINE=97251357; PubMed=9097039;
 RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
 RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
 RA Kitagawa M., Makino K., Miki T., Mizouchi K., Mori H., Mori T.,
 RA Motomura K., Nakano S., Nakamura Y., Nishimoto H., Nishio Y.,
 RA Oshima T., Saito N., Sampaio G., Seki Y., Sivasubraman S.,
 RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
 RA Yamamoto Y., Horikuchi T.;
 RT "A 570-kb DNA sequence of the *Escherichia coli* K-12 genome
 RT corresponding to the 28.0-40.1 min region on the linkage map.";
 RL DNA Res. 3:363-377(1996).
 RN [14]
 RP SEQUENCE OF 1-89 FROM N.A.
 RP MEDLINE=84221385; PubMed=6328431;
 RA Miles J.S., Guest J.R.;
 RT "Complete nucleotide sequence of the fumarase gene *fumA*, of
 RT *Escherichia coli*.";
 RL Nucleic Acids Res. 12:3631-3642(1984).
 RN [15]
 RP SEQUENCE OF 1-20.
 RP MEDLINE=92011457; PubMed=1917897;
 RA Ueda Y., Yumoto N., Tokushige M., Fukui K., Ohya-Nishiguchi H.;
 RT "Purification and characterization of two types of fumarase from
 RT *Escherichia coli*.";
 RL J. Biochem. 109:728-733(1991).
 RN [16]
 RP SEQUENCE OF 1-15, AND CHARACTERIZATION.
 RP MEDLINE=93267645; PubMed=8496960;
 RA Weaver T.M., Levitt D.G., Banaszak L.J.;
 RT "Purification and crystallization of fumarase C from *Escherichia*
 RT *coli*.";
 RL J. Mol. Biol. 231:141-144(1993).
 RN [17]
 RP BIOCHEMICAL ANALYSIS OF FUMA AND FUMC.
 RP MEDLINE=88193096; PubMed=3282546;
 RA Woods S.A., Schwartzbach S.D., Guest J.R.;
 RT "Two biochemically distinct classes of fumarase in *Escherichia coli*.";
 RL Biochim. Biophys. Acta 954:14-26(1988).
 RN [18]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
 RP MEDLINE=97065812; PubMed=8909293;
 RA Weaver T., Banaszak L.;
 RT "Crystallographic studies of the catalytic and a second site in
 RT fumarase C from *Escherichia coli*.";
 RL Biochemistry 35:13955-13965(1996).
 RN [19]
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
 RP MEDLINE=97253450; PubMed=9098893;
 RA Weaver T., Lees M., Banaszak L.;
 RT "Mutations of fumarase that distinguish between the active site and a
 RT nearby dicarboxylic acid binding site.";
 RL Protein Sci. 6:834-842(1997).
 CC -1- CATALYTIC ACTIVITY: (S)-malate = fumarate + H(2)O.
 CC -1- PATHWAY: Tricarboxylic acid cycle.
 CC -1- SUBUNIT: Homotetramer.

CC -1- SIMILARITY: TO OTHER THERMOSTABLE CLASS II FUMARASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 CC EMBL, X04065; CA427698.1; -;
 CC EMBL, AB000256; AAC74683.1; -;
 CC EMBL, D90803; BAA15349.1; -;
 CC EMBL, D90804; BAA15359.1; -;
 CC EMBL, D90805; BAA15363.1; -;
 CC EMBL, X00522; CA425205.1; -;
 CC PIR, S07138; UFEQ.
 CC PDB, 1FUO; 12-MAR-97.
 CC PDB, 1FUP; 12-MAR-97.
 CC PDB, 1FUQ; 12-MAR-97.
 CC PDB, 1FUR; 23-JUL-97.
 CC PDB, 2FUS; 23-JUL-97.
 CC PDB, 1KQ7; 23-AUG-02.
 CC ECO2DBASE; H050.4; 6TH EDITION.
 CC EcoGene; E010358; *fumC*.
 CC InterPro; IPR005677; *Fum hydII*.
 CC InterPro; IPR000362; Fumarate_lyase.
 CC Pfam; PF00206; *lyase 1, 1*.
 CC PRINTS; PR00149; FUMARATE_LYASE.
 CC TIGRFAMs; TIGR00979; *fumC II*; 1.
 CC PROSITE; PS00163; FUMARATE_LYASES, 1.
 CC KX *lyase*; Tricarboxylic acid cycle; 3D-structure; Complete proteome.
 CC ACT SITE 188 188
 CC FT BINDING 324 324
 CC FT STRAND 4 9
 CC FT TURN 10 11
 CC FT STRAND 12 17
 CC FT TURN 18 19
 CC FT HELIX 24 32
 CC FT STRAND 39 39
 CC FT HELIX 42 61
 CC FT TURN 62 63
 CC FT HELIX 67 81
 CC FT TURN 82 85
 CC FT HELIX 86 88
 CC FT STRAND 92 92
 CC FT TURN 97 98
 CC FT HELIX 100 117
 CC FT TURN 118 119
 CC FT TURN 124 125
 CC FT HELIX 130 134
 CC FT TURN 135 137
 CC FT HELIX 140 158
 CC FT TURN 159 159
 CC FT HELIX 160 177
 CC FT TURN 178 180
 CC FT STRAND 182 187
 CC FT TURN 188 189
 CC STRAND 190 196
 CC HELIX 197 222
 CC TURN 223 223
 CC STRAND 224 225
 CC TURN 227 228
 CC TURN 230 232
 CC TURN 236 237
 CC TURN 240 241
 CC STRAND 242 254
 CC STRAND 259 260
 CC HELIX 264 269
 CC HELIX 272 298
 CC TURN 299 299
 CC STRAND 309 310
 CC TURN 322 323

Db 16 SNVNNRNDLDDV 28

RESULT 7

PEP3_YEAST STANDARD; PRT; 918 AA.

ID P27801;

AC 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Vacuolar membrane protein PEP3.

GN PEP3 OR VPS18 OR YLR148W OR YL9634.2.

OS Saccharomyces cerevisiae (Baker's Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OK NCBI_TaxId=4932;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=92049305; PubMed=1944264;

RA Preston R., Manolson M.F., Becherer K., Weindhammer E.,

RA Kirpatrick D., Wright R., Jones E.W.,

RT "Isolation and characterization of PEP3, a gene required for vacuolar

RT biogenesis in *Saccharomyces cerevisiae*."

RL Mol. Cell. Biol. 11:5801-5812(1991).

RN (2)

RN SEQUENCE FROM N.A.

RX MEDLINE=92049306; PubMed=1840635;

RA Robinson J.S., Graham T.R., Bmr S.D.,

RT "A putative zinc finger protein, *Saccharomyces cerevisiae* Vps18p,

RT affects late golgi functions required for vacuolar protein sorting

RT and efficient alpha-factor promitome maturation."

RL Mol. Cell. Biol. 11:5813-5824(1991).

RN (3)

RP SEQUENCE FROM N.A.

RC STRAIN=8286C / AB972;

RX MEDLINE=97313267; PubMed=9169871;

RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansoorge W.,

RA Bens V., Bruckner M., Delius H., Dubois E., Duesterhoeft A.,

RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,

RA Heuss-Neitzel D., Hilbert H., Hilger F., Klose T., Moestl D.,

RA Louis E.J., Messenguy F., Nemes H.-W., Miosga T., Moestl D.,

RA Mueller-Auer S., Nentwich U., Obermayer B., Piravandi E., Pohl T.M.,

RA Portetle D., Punelle B., Reckmann S., Rieger M., Rinke M., Rose M.,

RA Scharte M., Scherens B., Scholler P., Schwager C., Schwarz S.,

RA Underwood A.P., Urestarazu L.A., Vandebol M., Verhaselt P.,

RA Viendeels F., Voet M., Volckaert G., Voss H., Wandt R., Wedler E.,

RA Wedler H., Zimmermann F.K., Zollner A., Hant J., Hehlsel J.D.,

RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XII."

RL Nature 387:87-90(1997).

CC -1- FUNCTION: REQUIRED FOR VACUOLAR BIOGENESIS.

CC -1- SUBCELLULAR LOCATION: CYTOSOLIC FACE OF THE VACUOLAR MEMBRANE.

CC -1- SIMILARITY: Contains 1 C3HC2-type zinc finger.

CC -1- SIMILARITY: Contains 1 C3HC2-type zinc finger.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL, M65144; AAA34852.1; -

DR EMBL, Z73320; CA937720.1; -

DR EMBL, U53879; AAB82382.1; -

DR PIR, A41943; A41943.

DR SGD, S0004138; PEP3.

DR GO, GO:0000306; C:cytosolic vacuolar membrane; IDA.

DR GO, GO:0000329; C:vacuolar membrane (sensu Fungi); IDA.

DR GO, GO:0005515; P:protein binding activity; IMP.

DR GO, GO:0006895; P:Golgi to endosome transport; IGI.

DR GO, GO:0042145; P:homotypic vacuole fusion (non-autophagic); IDA.

DR GO, GO:0045324; P:late endosome to vacuole transport; IGI.

DR GO, GO:0006904; P:non-selective vesicle docking; IMP.

DR InterPro: IPR00547; Clathrin_repeat.

DR InterPro: IPR001841; Znf_ring.

DR Pfam, PF06337; Clathrin_1.

DR Pfam, PF05131; PEP3_Vps18_1.

DR SMART, SM00299; CLH; 1.

DR SMART, SM00184; RING; 1.

DR Zinc-finger; Membrane.

FT ZN FING 826 851

FT C3HC2-TYPE.

SC SEQUENCE 918 AA, 107398 MW, 37AE3C34AE0470B8 CRC64;

Query Match 59.4%; Score 38; DB 1; Length 918;

Best Local Similarity 77.8%; Pred. No. 49;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 NMANNELSE 12

Db 786 NMKNNEISE 794

RESULT 8

RELX_HORSE STANDARD; PRT; 182 AA.

ID P22969; Q28907;

AC 01-AUG-1991 (Rel. 19, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Prerelaxin precursor (RXN).

GN RLN.

OS Equus caballus (Horse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

OK NCBI_TaxId=9796;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=Hokkaido; TISSUE=Placenta;

RA Min K., Shiota K., Ogawa T.,

RT "Molecular cloning of equine preprorelaxin cDNA."

RL J. Reprod. Dev. 42:171-178(1996).

RN (2)

RP SEQUENCE OF 32-174 FROM N.A.

RC TISSUE=Placenta;

RX MEDLINE=95359320; PubMed=7543295;

RA Klonisch T., Ryan P.L., Yamashiro S., Porter D.G.,

RT "Partial complementary deoxyribonucleic acid cloning of equine relaxin

RT messenger ribonucleic acid, and its localization within the equine

RT placenta."

RL Biol. Reprod. 52:1307-1315(1995).

RN (3)

RP SEQUENCE OF 26-53 AND 163-182.

RC TISSUE=Placenta;

RX MEDLINE=91275796; PubMed=2055195;

RA Stewart D.R., Nevins B., Hadas E., Vandlen R.,

RT "Affinity purification and sequence determination of equine relaxin."

RL Endocrinology 129:375-383(1991).

CC -1- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN

CC TO PRODUCE DILATATION OF THE BIRTH CANAL IN MANY MAMMALS.

CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO

CC DISULFIDE BONDS.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL, AB000201; BAA19069.1; -

DR EMBL, S78800; AAB35036.1; -

DR HSP; P01348; IRLX.
 DR InterPro: IPR004825; Ins/IGF/relex.
 DR Pfam: PF00069; Insulin; 1.
 DR SMART; SM00078; IIGF; 1.
 DR PROSITE; PS00262; INSULIN; 1.
 DR Insulin family; Hormone; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 53 RELAXIN B CHAIN.
 FT PROPEP 54 156 CONNECTING PEPTIDE.
 FT CHAIN 161 182 RELAXIN A CHAIN.
 FT DISULFID 35 169 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 47 182 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 168 173 BY SIMILARITY.
 FT CONFLICT 66 66 A -> V (IN REF. 2).
 FT CONFLICT 133 133 L -> Q (IN REF. 2).
 SQ SEQUENCE 182 AA; 20721 MW; E5C9414303A838B8 CRC64;
 Query Match 56.2%; Score 36; DB 1; Length 182;
 Best Local Similarity 54.5%; Pred. No. 18;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Qy 2 SDSNNMMNLESEV 12
 Db 120 DSNNLLEFEFE 130
 RESULT 9
 THII SULSO STANDARD; PRT; 371 AA.
 AC 098066;
 DT 28-FEB-2003 (Rel. 41, Last Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DN Probable thiamine biosynthesis protein thii.
 GN THII OR SSO0333.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 CC Sulfolobus.
 CC NCBI_TaxId=2287;
 CC (1)
 RN SEQUENCE FROM N.A.
 RP STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427276;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aweez M.J., Chan-Welner C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moore A., Fraus J., Fletcher C., Gordon P.M.K.,
 RA Helfkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Carreec R.A., Ragan M.A., Sensen C.W., Van der Oost J.,
 RA "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 CC -1- FUNCTION: Required for the synthesis of the thiazole moiety (by
 similarity).
 CC -1- PATHWAY: Thiamine biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: BELONGS TO THE THII FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE006668; AAK40666.1; -.
 DR PIR; C90176; C90176.
 DR HAMAP; MF_00021; -; 1.
 DR InterPro; IPR003720; THII.
 DR InterPro; IPR004114; THUMP_dom.
 DR Pfam; PF02568; Thii; 1.
 DR Pfam; PF02926; THUMP; 1.

DR TIGRFAMS; TIGR00342; TIGR00342; 1.
 KW Thiamine biosynthesis; Complete proteome.
 SQ SEQUENCE 371 AA; 41847 MW; 6D9F628AEE63F44 CRC64;
 Query Match 56.2%; Score 36; DB 1; Length 371;
 Best Local Similarity 53.8%; Pred. No. 40;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 SDSNNMMNLESEV 13
 Db 51 NDNMMNMLESEV 63
 RESULT 10
 ASPA_HELPJ STANDARD; PRT; 468 AA.
 ID ASPA_HELPJ
 AC 092115;
 DT 16-OCT-2001 (Rel. 40, Last Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Aspartate ammonia-lyase (EC 4.3.1.1) (Aspartase).
 GN ASPA OR JHP0594.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Helicobacteriaceae; Helicobacter.
 CC NCBI_TaxId=85963;
 CC (1)
 RN SEQUENCE FROM N.A.
 RP MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Dolg P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.D., Caruso A., Uria-Nickelsen M., Mills D.W., Ives C.,
 RA Gibson R., Werberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 gastric pathogen Helicobacter pylori.";
 RL Nature 397:176-180(1999).
 CC -1- CATALYTIC ACTIVITY: L-aspartate = fumarate + NH(3).
 CC -1- SUBUNIT: Homotetramer (by similarity).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE001491; AAD06167.1; -.
 DR PIR; H71913; H71913.
 DR HSP; P04422; IJWS.
 DR InterPro; IPR004708; ASPA.
 DR InterPro; IPR000362; Fumarate_Lyase.
 DR Pfam; PF00206; Lyase_1; 1.
 DR PRINTS; PR00149; FUMATELYASE.
 DR TIGRFAMS; TIGR00839; aspa; 1.
 DR PROSITE; PS00163; FUMARATE_LYASES; 1.
 KW Lyase; Complete proteome.
 SQ SEQUENCE 468 AA; 51937 MW; B7C6200CBA19442C CRC64;
 Query Match 56.2%; Score 36; DB 1; Length 468;
 Best Local Similarity 60.0%; Pred. No. 52;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 SDSNNMMNLE 10
 Db 99 TSTNNMMNEV 108
 RESULT 11
 ASPA_HELPJ STANDARD; PRT; 468 AA.
 ID ASPA_HELPJ
 AC P56159;

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Aspartate aminotransferase (EC 4.3.1.1) (Aspartase).
GN ASPA OR HP0649.
OC Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteriaceae; Helicobacter.
OX NCBI_TaxId=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.F., White O., Kervatage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgibbon L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RT Nature 388:539-547(1997).
CC -1- CATALYTIC ACTIVITY: L-aspartate = fumarate + NH(3).
CC -1- SUBUNIT: Homotetramer (By similarity).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL, AE000578; AAD07709.1; -.
CC PIR, A64601; A64601.
CC HSSP, P04422; 1J5W.
CC TIGR, HP0649; -.
DR InterPro, IPR004708; ApsA.
DR InterPro, IPR000362; Fumarate_lyase.
DR Pfam, PF00206; Lyase_1; 1.
DR PRINTS, PR00149; FUMARATELYASE.
DR TIGRFAMs, TIGR00839; aspa; 1.
DR PROSITE, PS00163; FUMARATE_LYASES; 1.
KW Lyase; Complete proteome.
SQ SEQUENCE 468 AA; 51981 MW; 60111869302836A CRC64;

Query Match 56.2%; Score 36; DB 1; Length 468;
Best Local Similarity 60.0%; Pred. No. 52;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SDSNNMMNEL 10
Db 99 TSTNNMNEV 108

RESULT 12
SECG BORBU STANDARD; PRT; 125 AA.
AC 051083;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable protein-export membrane protein secg.
GN SECG OR BB0054.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxId=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;

```

```

RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Castens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kervatage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Wathey L., McDonald L., Attlich P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RT Nature 390:580-586(1997).
CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT. PARTICIPATES IN AN EARLY
CC EVENT OF PROTEIN TRANSLLOCATION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SECG FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL, AE001119; AAC66457.1; -.
CC PIR, F70106; F70106.
DR TIGR, BB0054; -.
DR InterPro, IPR004692; Secg.
DR Pfam, PF03840; Secg; 1.
DR PRINTS, PR01651; SECGEXPORT.
DR TIGRFAMs, TIGR00810; secg; 1.
KW Protein transport; Translocation; Transmembrane; Complete proteome.
FT TRANSMEM 11 31
FT TRANSMEM 59 79
FT POTENTIAL.
SQ SEQUENCE 125 AA; 14015 MW; E0C2775AC19CE8F7 CRC64;

Query Match 54.7%; Score 35; DB 1; Length 125;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SDSNNMMNEL 12
Db 108 SEDANINNIKE 119

RESULT 13
VNSI HRSVA STANDARD; PRT; 139 AA.
AC P04544;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE Nonstructural protein 1 (Nonstructural protein 1C).
GN 1C OR NS1.
OS Human respiratory syncytial virus (strain A2).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxId=11259;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86045905; PubMed=2998021;
RA Collins P.L., Wertz G.W.;
RT "Nucleotide sequences of the 1B and 1C nonstructural protein mRNAs of
RT human respiratory syncytial virus."
RT Virology 143:442-451(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85237684; PubMed=4009789;
RA Elango N., Sakake M., Venkatesan S.;
RT "RNA sequence of three respiratory syncytial virus genes encoding
RT two nonstructural proteins and a 22k structural protein."
RT J. Virol. 55:101-110(1985).

```



```

RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE#95266253; PubMed#747420;
RA Connor M., Crowe J.E. Jr., Firestone C.Y., Murphy B.R., Collins P.L.;
RT "A cold-passaged, attenuated strain of human respiratory syncytial
RL virus contains mutations in the F and L genes.";
RN Virology 208:478-484(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE#97187925; PubMed#9035372;
RA Crowe J.E. Jr., Firestone C.Y., Whitehead S.S., Collins P.L.,
RA Murphy B.R.;
RT "Acquisition of the ts phenotype by a chemically mutagenized cold-
RT passaged human respiratory syncytial virus vaccine candidate results
RT from the acquisition of a single mutation in the polymerase (L)
RT gene.";
RL Virus Genes 13:269-273(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, M11486; AAB59850.1; -
DR EMBL, U50362; AAB86656.1; -
DR EMBL, U50363; AAB86668.1; -
DR EMBL, U39662; AAC57020.1; -
DR EMBL, U63644; AAC5562.1; -
DR PIR, A94336; MNNZIC.
DR InterPro; IPR005059; Pnuemo_NSI.
DR Pfam; PF03438; Pnuemo_NSI; 1.
DR ProDom; PD009407; Pnuemo_NSI; 1.
DR Nonstructural protein.
SQ SEQUENCE 139 AA; 15567 MW; EBCC69B09BC46BD CRC64;

Query Match 54.7%; Score 35; DB 1; Length 139;
Best Local Similarity 60.0%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY 1 SDSNNM--MNEISEV 13
DB 118 SDSTMTNWNQJSEL 132

RESULT 14
VNS1_HRSVL STANDARD; PRT; 139 AA.
ID VNS1_HRSVL STANDARD; PRT; 139 AA.
AC Q86306;
DT 15-JUL-1998 (Rel. 36; Created)
DT 15-JUL-1998 (Rel. 36; Last sequence update)
DT 15-JUL-1998 (Rel. 36; Last annotation update)
DE Nonstructural protein 1 (Nonstructural protein 1C).
GN 1C OR NS1.
OS Human respiratory syncytial virus (subgroup A / strain Long).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11260;
RN [1]
RP SEQUENCE FROM N.A.
RA Mazumder B., Dupuy L.C., McLean T., Barik S.;
RA Submitter B., Dupuy L.C., McLean T., Barik S.;
RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; U35030; AAA79091.1; -
DR InterPro; IPR005099; Pnuemo_NSI.
DR Pfam; PF03438; Pnuemo_NSI; 1.
DR ProDom; PD009407; Pnuemo_NSI; 1.
DR Nonstructural protein.
SQ SEQUENCE 139 AA; 15539 MW; EBD879B7FCA36BD CRC64;

Query Match 54.7%; Score 35; DB 1; Length 139;
Best Local Similarity 60.0%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY 1 SDSNNM--MNEISEV 13
DB 118 SDSTMTNWNQJSEL 132

RESULT 15
F801_SCHMA STANDARD; PRT; 238 AA.
ID F801_SCHMA STANDARD; PRT; 238 AA.
AC P16463;
DT 01-AUG-1990 (Rel. 15; Created)
DT 01-AUG-1990 (Rel. 15; Last sequence update)
DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE Female specific 800 protein (FS800).
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidae;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Puerto Rican;
RX MEDLINE#89181810; PubMed#2927441;
RA Reis W.G., Kums J., Blanton R., Davis A.H.;
RT "Localization and pattern of expression of a female specific mRNA in
RT Schistosoma mansoni.";
RL Mol. Biochem. Parasitol. 32:113-119(1989).
CC -1- FUNCTION: FS800 is likely to have some function in the production
CC or maintenance of the schistosome egg.
CC -1- DEVELOPMENTAL STAGE: Highest level only in mature worms, i.e.,
CC during egg production.
CC -1- MISCELLANEOUS: The two F800 proteins are read from two
CC overlapping reading frames.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03999; AAA29883.1; -
DT DOMAIN 42 45
FT POLY-ASN.
FT DOMAIN 72 77
FT POLY-ASP.
SQ SEQUENCE 238 AA; 27993 MW; FDC8B9F4181B7F1 CRC64;

Query Match 54.7%; Score 35; DB 1; Length 238;
Best Local Similarity 63.6%; Pred. No. 36;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 DSNMNMNEISE 12
DB 189 DSNININEKPE 199

```

Search completed: November 26, 2003, 12:31:18
 Job time : 8.26506 secs

~~This~~ Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:25:40 ; Search time 12.0602 Seconds
(without alignments)
103.662 Million cell updates/sec

Title: US-09-230-111c-13
Perfect score: 64
Sequence: 1 SDSNNMNNELSEV 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:1*
1: PIR1:1*
2: PIR2:1*
3: PIR3:1*
4: PIR4:1*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	67.2	1475	2 T33318	hypothetical prote
2	41	64.1	193	2 H82091	conserved hypothet
3	41	64.1	347	2 A85912	hypothetical prote
4	41	64.1	401	2 H91067	probable site spec
5	40	62.5	193	2 S70966	fimbrial protein c
6	39	60.9	317	2 H69135	protein kinase - M
7	39	60.9	462	1 UFBSC8	hydrolase
8	39	60.9	462	2 E83830	hydrolase
9	39	60.9	462	2 B81862	hydrolase
10	39	60.9	462	2 C81082	hydrolase
11	39	60.9	463	2 A81281	hydrolase
12	39	60.9	467	1 UPEC	hydrolase
13	39	60.9	467	2 B90918	hydrolase
14	39	60.9	467	2 B85167	hydrolase C (import
15	39	60.9	467	2 A10690	hydrolase C (import
16	38	59.4	663	2 S69566	glutamine transp
17	38	59.4	918	2 A41943	vacuolar membrane
18	37	57.8	269	2 A85049	probable transpos
19	37	57.8	533	2 T01864	hypothetical prote
20	37	57.8	589	2 AC2370	hypothetical prote
21	37	57.8	719	2 G81332	polysubunit lecti
22	37	57.8	1011	2 T13055	dynam associated
23	37	57.8	1094	2 T13053	dynam associated
24	37	57.8	2206	2 UC5280	voltage-dependent
25	36	56.2	143	2 T40866	relaxin B,C and A
26	36	56.2	311	2 S23693	erythrocyte membra
27	36	56.2	371	2 C90176	chitin biosynthes
28	36	56.2	379	2 A71674	hypothetical prote

30	36	56.2	461	2 E71672	hydrolase
31	36	56.2	463	2 D97826	hydrolase
32	36	56.2	465	2 A80275	hydrolase
33	36	56.2	468	2 A64601	aspartate ammonia-
34	36	56.2	468	2 H71913	aspartate ammonia-
35	36	56.2	521	2 T01923	hypothetical prote
36	36	56.2	615	2 G87678	ATP-dependent DNA
37	36	56.2	855	2 S47533	glucose-6-phosphat
38	36	56.2	910	2 S40259	glucose-6-phosphat
39	36	56.2	992	2 S54396	protein L precursor
40	36	56.2	1040	2 T08190	hypothetical prote
41	36	56.2	3097	2 T28635	glutamate synthase
42	35	54.7	106	2 A82870	chlorodoxin un589
43	35	54.7	125	2 F70106	hypothetical prote
44	35	54.7	139	1 MNZIC	nonstructural prot
45	35	54.7	223	2 T07098	catechol oxidase (

ALIGNMENTS

RESULT 1
T33318
hypothetical protein ZK1055.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T33318
R:Geisler, C.; Bradshaw, H.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid ZK1055.
A:Reference number: Z21321
A:Accession: T33318
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1475 <GEI>
A:Cross-references: EMBL:AF068721; PDB:AAIC19259.1; GSPDB:GN00023; CESP:ZK1055.1
A:Experimental source: strain Bristol N2; clone ZK1055
C:Genetics:
A:Gene: CESP:ZK1055.1
A:Map position: 5
A:introns: 17/3; 218/2; 385/3; 486/3; 1124/3; 1212/3; 1400/2

Query Match
Best Local Similarity 67.2%; Score 43; DB 2; Length 1475;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 2 DSNMNNELSEV 13
1057 DTNLMNDQSEM 1068

RESULT 2
H82091
conserved hypothetical protein VC2303 [imported] - Vibrio cholerae (strain N16961 serog:
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: H82091
R:Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.T.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bae, S.; Qin, H.; Dragol, I.; Sellere, P.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: H82091
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-193 <HEI>
A:Cross-references: GB:AE004301; GB:AE003652; NID:9965665; PDB:AAFS447.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2303
A:Map position: 1

Query Match 64.1%; Score 41; DB 2; Length 193;
 Best Local Similarity 75.0%; Pred. No. 5.9;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SDSNNMNNELSE 12
 |||||
 49 SKSNKNANLEISE 60

RESULT 3
 A:Accession: A85912
 Hypothetical protein Z3945 [imported] - Escherichia coli (strain O157:H7, substrain EDLg)
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: A85912
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobeck, B.J.; Davis, N.W.; Lim, A.; DiMantana, E.; Potamocuts, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: A85912
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-347 <STO>
 A:Cross-references: GB:AE005174; NID:G12517077; PIDN:AA657757.1; GSPDB:GN00145; UMGPR:Z39
 A:Experimental source: strain O157:H7, substrain EDLg33
 C:Genetics:
 A:Gene: Z3945

Query Match 64.1%; Score 41; DB 2; Length 347;
 Best Local Similarity 63.6%; Pred. No. 11;
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 DSNMNNNEISE 12
 |||||
 161 NSSLNNNEISE 171

RESULT 4
 H91067
 Probable site specific recombinase [imported] - Escherichia coli (strain O157:H7, substr
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C:Accession: H91067
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurukawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Ggsawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
 DNA Res. 8, 11-22, 2001
 A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: H91067
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-401 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA836935.1; PID:G13362983; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: EC83512

Query Match 64.1%; Score 41; DB 2; Length 401;
 Best Local Similarity 63.6%; Pred. No. 13;
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 DSNMNNNEISE 12
 |||||
 161 NSSLNNNEISE 171

RESULT 5
 S70966
 fimbrial protein class IV precursor - Escherichia coli plasmids
 M:Alternate names: bundle-forming pilin; bundlin precursor; pilin class IV precursor
 C:Species: Escherichia coli
 C>Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 08-Oct-1999

C:Accession: S70966; S28000; S31203; A60871; S24934
 R:Stone, K.D.; Zhang, H.Z.; Carlson, L.K.; Donenberg, M.S.
 Mol. Microbiol. 20, 325-337, 1996
 A>Title: A cluster of fourteen genes from enteropathogenic Escherichia coli is sufficien
 A:Reference number: S70966; MUID:96310370; PMID:8733231
 A:Accession: S70966
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-193 <STO>
 A:Cross-references: EMBL:Z68186; NID:G1122399; PIDN:CAA92326.1; PID:G213454; PID:G112240
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995
 R:Donenberg, M.S.; Gilron, J.A.; Natario, J.F.; Kaper, J.B.
 Mol. Microbiol. 6, 3427-3437, 1992
 A>Title: A plasmid-encoded type IV fimbrial gene of enteropathogenic Escherichia coli as
 A:Reference number: S28000; MUID:93133122; PMID:1362446
 A:Accession: S28000
 A:Molecule type: DNA
 A:Residues: 1-193 <DON>
 A:Cross-references: EMBL:Z12295; NID:G41061; PIDN:CAA78167.1; PID:G41062
 A:Experimental source: plasmid pMAR2
 R:Schel, I.; Puente, J.L.; Murray, W.J.; Vuopio-Varkila, J.; Schoolnik, G.K.
 Mol. Microbiol. 7, 563-575, 1993
 A>Title: Cloning and characterization of the bundle-forming pilin gene of enteropathogen
 A:Reference number: S31203; MUID:93211298; PMID:8096320
 A:Accession: S31203
 A:Molecule type: DNA
 A:Residues: 1-149, 'S', 151-193 <SCH>
 A:Cross-references: GB:U27184; EMBL:L07028; NID:G1314250; PIDN:AAC4040.1; PID:G1314251
 R:Gilron, J.A.; Ho, A.S.Y.; Schoolnik, G.K.
 Science 254, 710-713, 1991
 A>Title: An inducible bundle-forming pilus of enteropathogenic Escherichia coli.
 A:Reference number: A60871; MUID:92054556; PMID:1683004
 A:Accession: A60871
 A:Molecule type: protein
 A:Residues: 'W', 15-40, 'I', <GIR>
 A:Experimental source: strain B171; plasmid BAF
 C:Genetics:
 A:Gene: bfpA; bfp
 A:Gene: plasmid
 F:11-13/Domain: signal sequence #status predicted <SIG>
 F:14-193/Product: fimbrial protein class IV #status predicted <MAT>

Query Match 62.5%; Score 40; DB 2; Length 193;
 Best Local Similarity 61.5%; Pred. No. 8.9;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SDSNNMNNELSEV 13
 |||||
 41 SDSNKSQNAISEV 53

RESULT 6
 H69135
 protein kinase - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: H69135
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 Qiu, D.; Spadofora, R.; Vitale, R.; Wang, J.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
 K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Neelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: H69135
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-317 <MTH>
 A:Cross-references: GB:AE000814; GB:AE000666; NID:G2621334; PIDN:AAB84789.1; PID:G262133
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH283
 A:start codon: GTG

Query Match 60.9%; Score 39; DB 2; Length 317;
 Best Local Similarity 58.3%; Pred. No. 23;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SDSNNMNEV 13
 : |||||:
 : |||||:
 DB 82 DSTMNSNELOKI 93

RESULT 7
 UPGSC
 fumarate hydratase (EC 4.2.1.2) - *Bacillus subtilis*
 N/Alternate names: fumarase
 C/Species: *Bacillus subtilis*
 C/Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #ext_change 16-Jun-2000
 C/Accession: A23033; A33488; A69600
 R/Miles, J.S.; Guest, J.R.
 Nucleic Acids Res. 13, 131-140, 1985
 A/Title: Complete nucleotide sequence of the fumarate gene (cfc) of *Bacillus subtilis*
 A/Reference number: A23033; MUID:85215459; PMID:3923430
 A/Accession: A23033
 A/Molecule type: DNA
 A/Residues: 1-462 <N1L>
 A/Cross-references: GB:X01701; NID:G39842; PIDN:CAA25849.1; PID:G39844
 A/Experimental source: strain 168
 R/Price, V.A.; Peavers, I.M.; Molt, A.
 J. Bacteriol. 171, 5933-5939, 1989
 A/Title: Role of sigma H in expression of the fumarate gene (cfc) in vegetative cells
 A/Reference number: A33488; MUID:90036677; PMID:2509423
 A/Accession: A33488
 A/Status: not compared with conceptual translation
 A/Molecule type: DNA
 A/Residues: 1-50, 'N' <PRI>
 R/Kunet, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beren
 C.; Bron, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
 A.; Brillich, S.D.; Emerson, P.T.; Emtian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gall
 Koest, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holbappel, S.; Hosono, S.; Hullo, M.F.
 Leother, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudeg, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon
 A/Authors: Schleich, S.; Schreier, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 akuch, M.; Tanakoshi, A.; Tanaka, T.; Terpetra, A.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A/Reference number: A69580; MUID:98044033; PMID:9384377
 A/Accession: A69580
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-65, 'A', 67-462 <KUN>
 A/Cross-references: GB:Z99120; GB:AL009126; NID:G2635613; PIDN:CAB15294.1; PID:G2635801
 A/Experimental source: strain 168
 C/Comment: This is a class II fumarase, it is thermostable.
 C/Genetics:
 A/Gene: cfc
 A/Map position: 295 (degrees)
 C/Complex: homotetramer
 C/Function:
 A/Description: catalyzes the stereospecific interconversion of fumarate and L-malate
 A/Pathway: tricarboxylic acid cycle
 C/Superfamily: fumarate hydratase
 C/Keywords: ammonia-lyase; carbon-nitrogen lyase; carbon-oxygen lyase; homotetramer; hyd
 F,276/Active site: His #status predicted

Query Match 60.9%; Score 39; DB 1; Length 462;
 Best Local Similarity 70.0%; Pred. No. 34;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SDSNNMNEV 10
 : |||||:
 : |||||:

DB 99 TOSNNMNEV 108

RESULT 8
 E83830
 fumarate hydratase cfb [imported] - *Bacillus halodurans* (strain C-125)
 C/Species: *Bacillus halodurans*
 C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #ext_change 15-Jun-2001
 C/Accession: E83830
 R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Maeki, N.; Fujii, F.; Hir
 Nucleic Acids Res. 28, 4317-4331, 2000
 A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 A/Reference number: A83650; MUID:20512582; PMID:11058132
 A/Accession: E83830
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-462 <STO>
 A/Cross-references: GB:AP001512; GB:BA000004; NID:G10174030; PIDN:BA05164.1; GSPDB:GN00
 A/Experimental source: strain C-125
 C/Genetics:
 A/Gene: cfb
 C/Superfamily: fumarate hydratase

Query Match 60.9%; Score 39; DB 2; Length 462;
 Best Local Similarity 70.0%; Pred. No. 34;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SDSNNMNEV 10
 : |||||:
 : |||||:
 DB 99 TOSNNMNEV 108

RESULT 9
 B81862
 fumarate hydratase (EC 4.2.1.2) II NMA1670 [imported] - *Neisseria meningitidis* (strain
 C/Species: *Neisseria meningitidis*
 C/Date: 05-May-2000 #sequence_revision 05-May-2000 #ext_change 02-Feb-2001
 C/Accession: B81862
 R/Pearhill, J.; Achman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
 J.; Holtz, S.; Jagers, K.; Leather, S.; Mole, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A/Title: Complete genome sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
 A/Reference number: A81775; MUID:20222556; PMID:10761919
 A/Accession: B81862
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-462 <PAR>
 A/Cross-references: GB:AL162756; GB:AL157959; NID:G7380091; PIDN:CAB4898.1; PID:G738031
 A/Experimental source: serogroup A, strain Z2491
 C/Genetics:
 A/Gene: fumuC; NMA1670
 C/Superfamily: fumarate hydratase
 C/Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 60.9%; Score 39; DB 2; Length 462;
 Best Local Similarity 70.0%; Pred. No. 34;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SDSNNMNEV 10
 : |||||:
 : |||||:
 DB 99 TOSNNMNEV 108

RESULT 10
 C81082
 fumarate hydratase, class II NMB1458 [imported] - *Neisseria meningitidis* (strain MCS8 sc
 C/Species: *Neisseria meningitidis*
 C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #ext_change 19-Jan-2001
 C/Accession: C81082
 R/Petelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Bisen, J.J.
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
 Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizzo, M.
 Science 287, 1809-1815, 2000

A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V&
A/Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A/Reference number: A81000; PMID:2015755; PMID:10710307
A/Accession: C81082
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-462 <TEXT>
A/Cross-references: GB:AE002495; GB:AE002098; NID:G7226690; PIDN:AA41817.1; PID:G722669
A/Experimental source: serogroup B, strain MC58
C/Genetics:
A/Gene: NMB1458
C/Superfamily: fumarate hydratase

Query Match 60.9%; Score 39; DB 2; Length 462;
Best Local Similarity 70.0%; Pred. No. 34;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SDSNNMMNEL 10
: |||||:
Db 99 TOSNNMMNEV 108

RESULT 11
A81281
fumarate hydratase (EC 4.2.1.2) Cj1364c [imported] - *Campylobacter jejuni* (strain NCTC 1
C/Species: *Campylobacter jejuni*
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C/Accession: A81281
R/ParKhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A/Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hyf
A/Reference number: A81250; PMID:20150912; PMID:10688204
A/Accession: A81281
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-463 <PAR>
A/Cross-references: GB:AL139078; GB:AL111168; NID:G6968723; PIDN:CAB7291.1; PID:G696879
A/Experimental source: serotype O2, strain NCTC 11168
C/Genetics:
A/Gene: fumC; Cj1364c
C/Superfamily: fumarate hydratase
C/Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 60.9%; Score 39; DB 2; Length 463;
Best Local Similarity 70.0%; Pred. No. 34;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SDSNNMMNEL 10
: |||||:
Db 99 TOSNNMMNEV 108

RESULT 12
UPEC
fumarate hydratase (EC 4.2.1.2) fumC [validated] - *Escherichia coli* (strain K-12)
N/Alternate names: fumarase C; fumarate hydratase class II; protein g48
C/Species: *Escherichia coli*
C/Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 01-Mar-2002.
C/Accession: S07138; PS0181; S33825; E64917; A05151
R/Woods, S.A.; Miles, J.S.; Roberts, R.B.; Guest, J.R.
Biochem. J. 237, 547-557, 1986
A/Title: Structural and functional relationships between fumarase and aspartase. Nucleot
A/Reference number: S07138; PMID:87099873; PMID:3541901
A/Accession: S07138
A/Molecule type: DNA
A/Residues: 1-467 <WOO>
A/Cross-references: EMBL:X04065; NID:G41512; PIDN:CAA27698.1; PID:G41513
R/Ieda, Y.; Yumoto, N.; Tokushige, M.; Fukui, K.; Ohya-Nishiguchi, H.
J. Biochem. 109, 728-733, 1991
A/Title: Purification and characterization of two types of fumarase from *Escherichia coli*
A/Reference number: PX0048; PMID:92011457; PMID:1917897
A/Accession: PS0181

A/Molecule type: protein
A/Residues: 1-20 <U&D>
R/Guest, J.R.; Miles, J.S.; Roberts, R.B.; Woods, S.A.
J. Gen. Microbiol. 131, 2971-2984, 1985
A/Title: The fumarase genes of *Escherichia coli*: location of the fumb gene and discovery
A/Reference number: A92783; PMID:86142617; PMID:3005475
A/Contents: annotation; identification of structural gene
R/Weaver, T.W.; Levitt, D.G.; Banaszak, L.J.
J. Mol. Biol. 231, 141-144, 1993
A/Title: Purification and crystallization of fumarase C from *Escherichia coli*.
A/Reference number: S33825; PMID:93267645; PMID:8496960
A/Accession: S33825
A/Molecule type: protein
A/Residues: 1-15 <MEA>
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC
A.; Rose, D.U.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of *Escherichia coli* K-12.
A/Reference number: A64720; PMID:97426617; PMID:9278503
A/Accession: E64917
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-467 <BLAT>
A/Cross-references: GB:AE000256; GB:U00096; NID:G1787888; PIDN:AACT4683.1; PID:G1787896,
A/Experimental source: strain K-12, substrain MG1655
C/Genetics:
A/Gene: fumC
A/Map position: 35.5 min
A/Note: the gene coding for this protein was originally called g48 and was identified as
C/Complex: homotetramer [validated, PMID:93267645]
C/Function:
A/Description: EC 4.2.1.2 [validated, PMID:93267645]; fumarate hydratase; catalyzes the
A/Pathway: tricarboxylic acid cycle
A/Note: In *E. coli* three different fumarate hydratase genes (fuma, fumb, and fumC) have
A/Note: fumC-encoded fumarate hydratase is a thermostable and iron-independent class II
C/Superfamily: fumarate hydratase
C/Keywords: carbon-oxygen lyase; heat-stable protein; homotetramer; hydro-lyase; tricar
F:278/Active site: His #status predicted

Query Match 60.9%; Score 39; DB 1; Length 467;
Best Local Similarity 70.0%; Pred. No. 35;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SDSNNMMNEL 10
: |||||:
Db 100 TOSNNMMNEV 109

RESULT 13
E90918
fumarase C [imported] - *Escherichia coli* (strain O157:H7, substrain RIMD 0509952)
C/Species: *Escherichia coli*
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C/Accession: E90918
R/Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G
gasaawa, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gen
A/Reference number: A99629; PMID:21156231; PMID:11258796
A/Accession: E90918
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-467 <HAY>
A/Cross-references: GB:BA000007; PIDN:BAB35740.1; PID:G13361784; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain RIMD 0509952
C/Genetics:
A/Gene: EC92317
C/Superfamily: fumarate hydratase

Query Match 60.9%; Score 39; DB 2; Length 467;
Best Local Similarity 70.0%; Pred. No. 35;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SDSNNMNNEL 10
: |||||:
DB 100 TQSNMNNNEV 109

RESULT 14

B85767
fumarate C [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001
C/Accession: B85767
R/Perna, N.T.; Plunkett II, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Llin, A.; Dimalanta, E.; Potamoukis, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: AB5480; MUID:21074935; PMID:11206551
A/Accession: B85767
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-467 <STO>
A/Cross-References: GB:AB005174; NID:g12515592; PIDN:AA056598.1; GSPDB:GN00145; UMGF:226
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Gene: funcC
C/Superfamily: fumarate hydratase

Query Match 60.9%; Score 39; DB 2; Length 467;
Best Local Similarity 70.0%; Pred. No. 35;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SDSNNMNNEL 10
: |||||:
DB 100 TQSNMNNNEV 109

RESULT 15

AT0690
fumarate hydratase class II [imported] - Salmonella enterica subsp. enterica serovar Tyf
C/Species: Salmonella enterica subsp. enterica serovar Typh
A/Note: This species has also been called Salmonella typh
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C/Accession: AT0690
R.Parthill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A/Reference number: AB0502; MUID:21534947; PMID:11677608
A/Accession: AT0690
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-467 <PAR>
A/Cross-References: GB:AL513382; PIDN:CAD01898.1; PID:g16502742; GSPDB:GN00176
C/Genetics:
A/Gene: STY1653
C/Superfamily: fumarate hydratase

Query Match 60.9%; Score 39; DB 2; Length 467;
Best Local Similarity 70.0%; Pred. No. 35;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SDSNNMNNEL 10
: |||||:
DB 100 TQSNMNNNEV 109

Search completed: November 26, 2003, 12:36:03
Job time: 14.0602 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 26, 2003, 12:25:10 / Search time 28.3494 Seconds

(without alignments)
118.334 Million cell updates/sec

Title: US-09-230-111c-13

Perfect score: 64

Sequence: 1 SDSNNMMNELSEV 13

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

SPREMBL.23.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	67.2	1475	5	076447
2	41	64.1	193	16	09KPR3
3	41	64.1	347	16	08X402
4	41	64.1	401	16	08X3D8
5	40	62.5	170	2	093MA2
6	40	62.5	193	2	09F5U8
7	40	62.5	193	2	09J3W6
8	40	62.5	193	2	09ERG4
9	40	62.5	195	2	09F5V1
10	40	62.5	195	2	09F5U7
11	39	60.9	317	17	026383
12	39	60.9	461	16	08XQEB
13	39	60.9	462	16	09KCM4
14	39	60.9	462	16	09JYR9
15	39	60.9	462	16	09JYR0
16	39	60.9	467	16	08Z6R6

17	39	60.9	467	16	08ZPL7	08ZPL7 salmoneilla
18	39	60.9	467	16	08X769	08X769 escherichia
19	39	60.9	467	16	08FHA7	08FHA7 escherichia
20	39	60.9	490	16	08G7D1	08G7D1 bifidobacte
21	38	59.4	265	16	08EAB8	08EAB8 shewanella
22	38	59.4	326	12	09IV20	09IV20 human rotav
23	38	59.4	1145	5	09U0J0	09U0J0 plasmodium
24	38	59.4	2940	5	08IHP9	08IHP9 plasmodium
25	38	59.4	3279	5	09N4B9	09N4B9 caenorhabdi
26	38	59.4	3648	5	08IDR0	08IDR0 plasmodium
27	37.5	58.6	3633	5	08IHL0	08IHL0 dictyostel
28	37	57.8	127	2	09WTAS	09WTAS escherichia
29	37	57.8	139	5	08IDC1	08IDC1 plasmodium
30	37	57.8	178	16	08XKC2	08XKC2 clostridium
31	37	57.8	193	2	09F5U9	09F5U9 escherichia
32	37	57.8	193	2	08RNV3	08RNV3 escherichia
33	37	57.8	193	2	09ERTX0	09ERTX0 escherichia
34	37	57.8	195	2	09F5V0	09F5V0 escherichia
35	37	57.8	269	10	09M104	09M104 arabidopsis
36	37	57.8	378	10	08S862	08S862 oryza sativ
37	37	57.8	466	2	053246	053246 ticketelia
38	37	57.8	466	2	053251	053251 ticketelia
39	37	57.8	521	16	08DFB6	08DFB6 vibrio vuln
40	37	57.8	533	10	081510	081510 arabidopsis
41	37	57.8	588	10	09JIM6	09JIM6 arabidopsis
42	37	57.8	589	16	08YNP9	08YNP9 anabaena sp
43	37	57.8	628	5	097243	097243 plasmodium
44	37	57.8	719	16	09PN44	09PN44 campylobact
45	37	57.8	885	5	08IEM9	08IEM9 plasmodium

ALIGNMENTS

RESULT 1

076447 PRELIMINARY; PRT; 1475 AA.

AC 076447; 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

OS Hypothetical 166.8 KDa protein.

GN ZK1055.1.

OC Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;

OC Rhabditidae; Pelodexinae; Caenorhabditis.

OX NCBI_Taxid=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RX MEDLINE=99069613; PubMed=9851916;

RA None;

RT "Genome sequence of the nematode C. elegans: a platform for

investigating biology. The C. elegans Sequencing Consortium.";

RL Science 282:2012-2018(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RX "The sequence of C. elegans cosmid ZK1055.";

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RX "Direct Submission.";

RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

RL EMBL; AF068721; AAC19259.1;

DR WormPep; ZK1055.1; CE18470.

KW Hypothetical protein.

SQ SEQUENCE 1475 AA; 166762 MW; 0483113C3D5D8769 CRC64;

Query Match 67.2%; Score 43; DB 5; Length 1475;

Best Local Similarity 58.3%; Pred. No. 77;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 2 DSNMNNELISEV 13
: : : : :
DB 1057 DTNNMDQLSEM 1068

RESULT 2

Q9KPR3 PRELIMINARY; PRT; 193 AA.
AC Q9KPR3;
DT 01-OCT-2000 (TRENBLREL. 15, Created)
DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
DE Hypothetical protein VC2303.
GN VC2303.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_Taxid=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Drygiol I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Newman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae."
RL Nature 406:477-483 (2000).
DR EMBL; AE004301; AAF95447.1; -.
DR TIGR; VC2303; -.
DR InterPro; IPR003111; ION.
DR Pfam; PF02190; ION; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 193 AA; 22251 MW; D7CB416078B71A15 CRC64;

Query Match 64.1%; Score 41; DB 16; Length 193;
Best Local Similarity 75.0%; Pred. No. 24;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SPSNNMNEISE 12
: : : : :
DB 49 SKSNKNANLISE 60

RESULT 3

Q8X402 PRELIMINARY; PRT; 347 AA.
AC Q8X402;
DT 01-MAR-2002 (TRENBLREL. 20, Created)
DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Hypothetical protein Z3945.
GN Z3945.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_Taxid=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grodbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,

RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533 (2001).
DR EMBL; AE005493; AAG57757.1; -.
DR InterPro; IPR002104; Phage_Integrase.
DR Pfam; PF00589; Phage_Integrase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 347 AA; 40506 MW; 3E5250026FE2E5963 CRC64;

Query Match 64.1%; Score 41; DB 16; Length 347;
Best Local Similarity 63.6%; Pred. No. 42;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 DSNMNNELISE 12
: : : : :
DB 161 NSSLNNELISE 171

RESULT 4

Q8X3D8 PRELIMINARY; PRT; 401 AA.
AC Q8X3D8;
DT 01-MAR-2002 (TRENBLREL. 20, Created)
DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Putative site specific recombinase.
GN ECS3512.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_Taxid=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 050952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22 (2001).
DR EMBL; AP002562; BAB36935.1; -.
DR InterPro; IPR002104; Phage_Integrase.
DR Pfam; PF00589; Phage_Integrase; 1.
SQ SEQUENCE 401 AA; 76799 MW; 8C51A92F4B4E535 CRC64;

Query Match 64.1%; Score 41; DB 16; Length 401;
Best Local Similarity 63.6%; Pred. No. 48;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 DSNMNNELISE 12
: : : : :
DB 161 NSSLNNELISE 171

RESULT 5

Q93NA2 PRELIMINARY; PRT; 170 AA.
AC Q93NA2;
DT 01-DEC-2001 (TRENBLREL. 19, Created)
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Buntlin alpha (Fragment).
GN BFA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_Taxid=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=sg049692;
RA Mansfield K.G., Lin K.-C., Xia D., Schauer D., Mackey J., Lackner A.,

RA Carville A.;
 RT "Enteropathogenic Escherichia coli and ulcerative colitis in cotton-
 RT top tamarins (Saguinus oedipus).";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF382948; AAK6135.1; -
 RT NON TER
 FT 1
 SQ SEQUENCE 170 AA; 17681 MW; F0E49BCB1F0DECSF CRC64;

Query Match 62.5%; Score 40; DB 2; Length 170;
 Best Local Similarity 61.5%; Pred. No. 32;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SDSNNMMNELSEV 13
 |||| : ||||
 Db 18 SDSNKSQNAISEV 30

RESULT 6

Q9FSU8 PRELIMINARY; PRT; 193 AA.

AC Q9FSU8, 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE Bundlin alpha-1.

GN BPPA.

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RP [1]

RC SEQUENCE FROM N.A.

RA STRAIN=HSP19/4;

RL MEDLINE=20536453; PubMed=11083828;

RT Blank T.E., Zhong H., Bell A.L., Whitlam T.S., Donnenberg M.S.;

RT "Molecular Variation among Type IV Pilin (bfpA) Genes from Diverse

RT Enteropathogenic Escherichia coli Strains.";

RL Infect. Immun. 68:7028-7038(2000).

DR EMBL; AF304481; AAG16269.1; -

SQ SEQUENCE 193 AA; 20306 MW; F22414697FD20969 CRC64;

Query Match 62.5%; Score 40; DB 2; Length 193;
 Best Local Similarity 61.5%; Pred. No. 36;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SDSNNMMNELSEV 13
 |||| : ||||
 Db 41 SDSNKSQNAISEV 53

RESULT 7

Q933W6 PRELIMINARY; PRT; 193 AA.

AC Q933W6, 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Bundlin alpha-2.

GN BPPA.

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RP [1]

RC SEQUENCE FROM N.A.

RA STRAIN=Stoke W, 009-2710820, 2309-77, DIF043256, and RM191/1;

RX MEDLINE=20536453; PubMed=11083828;

RT Blank T.E., Zhong H., Bell A.L., Whitlam T.S., Donnenberg M.S.;

RT "Molecular Variation among Type IV Pilin (bfpA) Genes from Diverse

RT Enteropathogenic Escherichia coli Strains.";

RL Infect. Immun. 68:7028-7038(2000).

DE Bundlin beta-1.

GN BPPA.

RP SEQUENCE FROM N.A.

RC PLASWID=PB171;
 RX MEDLINE=99426847; PubMed=10496929;
 RA ToBe T., Hayashi T., Han C.G., Schoolnik G.K., Ohtsuo E.,
 RA Sasakawa C.;

RT "Complete DNA sequence and structural analysis of the enteropathogenic
 RT Escherichia coli adherence factor plasmid.";

RL Infect. Immun. 67:5455-5462(1999).

DR EMBL; AF304485; AAG16272.1; -

DR EMBL; AF304469; AAG16257.1; -

DR EMBL; AF304473; AAG16261.1; -

DR EMBL; AF304477; AAG16265.1; -

DR EMBL; AF304483; AAG16271.1; -

DR EMBL; AB024946; BAA84838.1; -

DR InterPro; IPR001120; Prok_N_methyln.

DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.

SQ SEQUENCE 193 AA; 20300 MW; 2FCE03E415FF026 CRC64;

Query Match 62.5%; Score 40; DB 2; Length 193;
 Best Local Similarity 61.5%; Pred. No. 36;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SDSNNMMNELSEV 13
 |||| : ||||
 Db 41 SDSNKSQNAISEV 53

RESULT 8

Q9ETG4 PRELIMINARY; PRT; 193 AA.

AC Q9ETG4, 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Bundlin alpha-3.

GN BPPA.

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RP [1]

RC SEQUENCE FROM N.A.

RA STRAIN=RN410/1, 10, 010-311082, and 0659-79;

RX MEDLINE=20536453; PubMed=11083828;

RT Blank T.E., Zhong H., Bell A.L., Whitlam T.S., Donnenberg M.S.;

RT "Molecular Variation among Type IV Pilin (bfpA) Genes from Diverse

RT Enteropathogenic Escherichia coli Strains.";

RL Infect. Immun. 68:7028-7038(2000).

DR EMBL; AF304484; AAG16272.1; -

DR EMBL; AF304468; AAG16256.1; -

DR EMBL; AF304470; AAG16258.1; -

DR EMBL; AF304472; AAG16260.1; -

DR InterPro; IPR001120; Prok_N_methyln.

DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.

SQ SEQUENCE 193 AA; 20227 MW; D727852C0528A83E CRC64;

Query Match 62.5%; Score 40; DB 2; Length 193;
 Best Local Similarity 61.5%; Pred. No. 36;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SDSNNMMNELSEV 13
 |||| : ||||
 Db 41 SDSNKSQNAISEV 53

RESULT 9

Q9FSV1 PRELIMINARY; PRT; 195 AA.

AC Q9FSV1, 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Bundlin beta-1.

GN BPPA.

```

OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=562;

RN
  [1]
  SEQUENCE FROM N.A.
  STRAIN=012-050982;
  MEDLINE=20536453; PubMed=11083828;
  RA Black T.E., Zhong H., Bell A.L., Whitlam T.S., Donnenberg M.S.;
  RT "Molecular Variation among Type IV Pilin (bfpA) Genes from Diverse
  Enteropathogenic Escherichia coli Strains.";
  RL Infect. Immun. 68:7028-7038(2000).
  DR EMBL; AF304471; AAG16259.1; -.
  DR InterPro; IPR001120; Prok N methyln.
  DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
  SQ SEQUENCE 195 AA; 20325 MW; BD848F6FE86DF7B5 CRC64;

Query Match
  Best Local Similarity 62.5%; Score 40; DB 2; Length 195;
  Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SDSNNMMNELSEV 13
  ||| : |||
  Db 41 SDSNKSQNAISEV 53

RESULT 10
QPSU7 PRELIMINARY; PRT; 195 AA.
ID Q9FSU7;
AC Q9FSU7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Bundlin beta-2.
GN BfPA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=562;
RN
  [1]
  SEQUENCE FROM N.A.
  STRAIN=2188-93;
  MEDLINE=20536453; PubMed=11083828;
  RA Black T.E., Zhong H., Bell A.L., Whitlam T.S., Donnenberg M.S.;
  RT "Molecular Variation among Type IV Pilin (bfpA) Genes from Diverse
  Enteropathogenic Escherichia coli Strains.";
  RL Infect. Immun. 68:7028-7038(2000).
  DR EMBL; AF304486; AAG16274.1; -.
  DR InterPro; IPR001120; Prok N methyln.
  DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
  SQ SEQUENCE 195 AA; 20307 MW; 8BC4C71310079F12 CRC64;

Query Match
  Best Local Similarity 62.5%; Score 40; DB 2; Length 195;
  Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SDSNNMMNELSEV 13
  ||| : |||
  Db 41 SDSNKSQNAISEV 53

RESULT 11
ID Q26383 PRELIMINARY; PRT; 317 AA.
AC Q26383;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Protein kinase.
GN MTH283.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.

```

```

OX NCBI_TaxId=187420;
RN
  [1]
  SEQUENCE FROM N.A.
  STRAIN-Delta H;
  MEDLINE=98037514; PubMed=9371463;
  RA Smith D.R., Doucette-Stamm U.A., Deloughery C., Lee H.-M., Dubois J.,
  RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
  RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
  RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
  RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
  RA McDougall S., Shiner G., Goyal A., Pietrovski S., Church G.M.,
  RA Daniels C.J., Mao J.-I., Rice P., Noelling U., Reeve J.N.;
  RT "Complete genome sequence of Methanobacterium thermoautotrophicum
  deltaH: functional analysis and comparative genomics.";
  RL J. Bacteriol. 179:7135-7155(1997).
  DR EMBL; AE000814; AAB84789.1; -.
  KW Complete proteome.
  SQ SEQUENCE 317 AA; 36630 MW; C71AF76D72A658B9 CRC64;

Query Match
  Best Local Similarity 60.9%; Score 39; DB 17; Length 317;
  Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 DSNMMNELSEV 13
  ||| : |||
  Db 82 DSTMNSNELOKI 93

RESULT 12
Q8XOE8 PRELIMINARY; PRT; 461 AA.
ID Q8XOE8;
AC Q8XOE8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Probable fumarate hydratase class II (Fumarate) protein
  (EC 4.2.1.2).
GN FUMC OR RSP1279 OR RS05324.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC plasmid megaplasmid.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Ralstoniaceae; Ralstonia.
OX NCBI_TaxId=305;
RN
  [1]
  SEQUENCE FROM N.A.
  STRAIN-SM11000;
  MEDLINE=21681879; PubMed=11823852;
  RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
  RA Arlat M., Billault A., Brotlier P., Camus J.C., Catolico L.,
  RA Chandler M., Cholet N., Claudel-Renard C., Cunnac S., Demange N.,
  RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
  RA Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
  RA Weisenbach J., Boucher C.A.;
  RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
  RL Nature 415:497-502(2002).
  DR EMBL; AL646083; CAD18430.1; -.
  DR InterPro; IPR000362; Fumarate lyase.
  DR InterPro; IPR005677; Fum hydII.
  DR Pfam; PF00206; Lyase_1; I.
  DR PRINTS; PR00149; FUMRATLYASE.
  DR TIGRFAMs; TIGR00979; fumC II; 1.
  DR PROSITE; PS00163; FUMARATE LYASES; 1.
  KW Lyase; Plasmid; Complete proteome.
  SQ SEQUENCE 461 AA; 49403 MW; 8A44D6B53DE18B CRC64;

Query Match
  Best Local Similarity 60.9%; Score 39; DB 16; Length 461;
  Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SDSNNMMNEL 10
  : ||| :
  Db 99 TOSNNMMNEV 108

```

RESULT 13

Q9CKX4 PRELIMINARY; PRT; 462 AA.
 AC Q9CKX4; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Fumarate hydratase.
 GN C17B OR BH145.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL; AP001512; BAB05164.1; -
 DR HSSP; P05042; 1P00.
 DR InterPro; IPR000362; Fumarate_lyase.
 DR InterPro; IPR005677; Fum_hydlt.
 DR Pfam; PF00206; lyase_1; 1.
 DR PRINTS; PRO0149; FUMARATELYASE.
 DR TIGRFAMs; TIGR00979; func_II; 1.
 DR PROSITE; PS00163; FUMARATE_LYASES; 1.
 KW Complete proteome.
 SQ SEQUENCE 462 AA; 50785 MW; 72601B22909B4899 CRC64;

Query Match

Best Local Similarity 70.0%; Score 39; DB 16; Length 462;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SDSNNMNNEL 10
 Db 99 TQSNMNNMNEV 108

RESULT 14

Q9JYR9 PRELIMINARY; PRT; 462 AA.
 AC Q9JYR9; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Fumarate hydratase, class II.
 GN NMB1458.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Ciftone H., Clark E.B.,
 RA Cotton M.D., Uitterback T.R., Knouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Maignant V., Pizza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappelli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 RT MC58.";
 RL Science 287:1809-1815(2000).
 DR EMBL; AE002495; AAP41817.1; -
 DR HSSP; P08417; 1YFM.

DR TIGR; NMB1458; -
 DR InterPro; IPR000362; Fumarate_lyase.
 DR InterPro; IPR005677; Fum_hydlt.
 DR Pfam; PF00206; lyase_1; 1.
 DR PRINTS; PRO0149; FUMARATELYASE.
 DR TIGRFAMs; TIGR00979; func_II; 1.
 DR PROSITE; PS00163; FUMARATE_LYASES; 1.
 KW Complete proteome.
 SQ SEQUENCE 462 AA; 49332 MW; 4B44ADD0D3489F250 CRC64;

Query Match
 Best Local Similarity 70.0%; Score 39; DB 16; Length 462;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SDSNNMNNEL 10
 Db 99 TQSNMNNMNEV 108

RESULT 15

Q9JTR0 PRELIMINARY; PRT; 462 AA.
 AC Q9JTR0; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Fumarate hydratase class II (EC 4.2.1.2).
 GN FUMC OR NMA1670.
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=65699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Z2491 / Serotype 4A;
 RX MEDLINE=2022556; PubMed=10761919;
 RA Parthill J., Achman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holtroyd S.,
 RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria
 RT meningitidis Z2491.";
 RL Nature 404:502-506(2000).
 DR EMBL; AL162756; CAB84898.1; -
 DR HSSP; P05042; 1P00.
 DR InterPro; IPR000362; Fumarate_lyase.
 DR InterPro; IPR005677; Fum_hydlt.
 DR Pfam; PF00206; lyase_1; 1.
 DR PRINTS; PRO0149; FUMARATELYASE.
 DR TIGRFAMs; TIGR00979; func_II; 1.
 DR PROSITE; PS00163; FUMARATE_LYASES; 1.
 KW lyase; Complete proteome.
 SQ SEQUENCE 462 AA; 49391 MW; 6F63BD506E26C85 CRC64;

Query Match
 Best Local Similarity 70.0%; Score 39; DB 16; Length 462;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SDSNNMNNEL 10
 Db 99 TQSNMNNMNEV 108

Search completed: November 26, 2003, 12:34:35
 Job time : 31.3494 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: November 26, 2003, 12:26:20 ; Search time 27.1084 Seconds

(Without alignments)
102.059 Million cell updates/sec

Title: US-09-230-111c-14

Perfect score: 79

Sequence: 1 PPTCGANGCRISTL 15

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 673684 seqs, 184443283 residues

Total number of hits satisfying chosen parameters: 673684

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications AA.*

1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	15	8	US-08-681-219-16
2	79	100.0	15	12	US-10-092-138-14
3	79	100.0	15	12	US-09-230-111c-14
4	41	51.9	77	15	US-10-106-698-4719
5	41	51.9	133	10	US-09-764-868-1068
6	41	51.9	133	11	US-09-955-999-91
7	40	50.6	126	15	US-10-106-698-6252
8	40	50.6	404	9	US-09-925-302-768
9	39	49.4	80	10	US-09-925-300-1118
10	39	49.4	161	15	US-10-156-761-1117
11	39	49.4	359	9	US-09-888-615-114
12	39	49.4	401	10	US-09-771-161A-145
13	39	49.4	780	10	US-09-771-161A-144
14	39	49.4	942	10	US-09-771-161A-235
15	39	49.4	942	10	US-09-771-161A-236

16	39	49.4	942	12	US-10-354-358-74	Sequence 74, Appl
17	38	48.1	92	11	US-09-755-108-3	Sequence 3, Appl1
18	38	48.1	94	15	US-10-106-698-5724	Sequence 5724, Ap
19	38	48.1	123	15	US-10-147-951B-2	Sequence 2, Appl1
20	38	48.1	168	12	US-10-135-545A-35	Sequence 35, Appl
21	38	48.1	223	14	US-10-005-168-2	Sequence 2, Appl1
22	38	48.1	244	9	US-09-815-242-10119	Sequence 10119, A
23	38	48.1	335	15	US-10-193-656-6	Sequence 6, Appl
24	38	48.1	335	15	US-10-177-293-362	Sequence 362, Appl
25	38	48.1	348	16	US-10-080-110-399	Sequence 399, App
26	38	48.1	386	15	US-10-147-951B-9	Sequence 9, Appl1
27	38	48.1	586	12	US-10-308-279-50	Sequence 50, Appl
28	38	48.1	1076	15	US-10-116-949-6	Sequence 6, Appl1
29	38	48.1	1145	15	US-10-116-949-2	Sequence 2, Appl1
30	38	48.1	1145	15	US-10-116-949-4	Sequence 4, Appl1
31	37	46.8	123	15	US-10-106-698-5744	Sequence 5744, Ap
32	37	46.8	310	15	US-10-023-597-64	Sequence 64, Appl
33	37	46.8	495	10	US-09-971-536-52	Sequence 52, Appl
34	37	46.8	502	15	US-10-146-473-52	Sequence 52, Appl
35	37	46.8	871	15	US-10-307-019-7	Sequence 7, Appl1
36	37	46.8	953	9	US-09-888-615-66	Sequence 66, Appl
37	37	46.8	1009	15	US-10-290-544-2	Sequence 2, Appl1
38	37	46.8	1009	15	US-10-290-544-5	Sequence 5, Appl1
39	37	46.8	1351	15	US-10-307-019-1	Sequence 1, Appl1
40	37	46.8	1510	15	US-10-307-019-4	Sequence 4, Appl1
41	37	46.8	1665	9	US-09-858-664A-2	Sequence 2, Appl1
42	37	46.8	2596	15	US-10-307-019-6	Sequence 6, Appl1
43	37	46.8	2630	14	US-10-077-130-2	Sequence 2, Appl1
44	37	46.8	7968	14	US-10-077-130-5	Sequence 5, Appl1
45	36.5	46.2	458	11	US-09-770-564-9	Sequence 9, Appl1

ALIGNMENTS

RESULT 1
US-08-681-219-16
Sequence 16, Application US/08681219
; Publication No. US20020058607A1
; GENERAL INFORMATION:
APPLICANT: Takaaki Sato and Junn Yanagisawa
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF
TITLE OF INVENTION: (PDI/DHR) DOMAIN AND USES THEREOF
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/681,219
FILING DATE: 22-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28, 678
REFERENCE/DOCKET NUMBER: 0575/48962/JPM/JCM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-681-219-16

Query Match 100.0%; Score 79; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPTCSQANSGRISTL 15
|||||

DB 1 PPTCSQANSGRISTL 15

RESULT 2
US-10-092-138-14

Sequence 14, Application US/10092138
Publication No. US20030170723A1
GENERAL INFORMATION:

APPLICANT: Sato, Taka-Aki

TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
TITLE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION

FILE REFERENCE: 65823/JPW/PT

CURRENT APPLICATION NUMBER: US/10/092,138

NUMBER OF SEQ ID NOS: 33

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 14

LENGTH: 15

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial

US-10-092-138-14

Query Match 100.0%; Score 79; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPTCSQANSGRISTL 15
|||||

DB 1 PPTCSQANSGRISTL 15

RESULT 3

US-09-230-111C-14

Sequence 14, Application US/09230111C

Publication No. US20030203414A1

GENERAL INFORMATION:

APPLICANT: Sato, Taka-Aki

APPLICANT: Yanagisawa, Junn

TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN

TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF (PDZ/DHR)

TITLE OF INVENTION: DOMAIN AND USES THEREOF

FILE REFERENCE: 48962-A-PCT-US

CURRENT APPLICATION NUMBER: US/09/230,111C

NUMBER OF SEQ ID NOS: 33

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 14

LENGTH: 15

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial

US-09-230-111C-14

Query Match 100.0%; Score 79; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPTCSQANSGRISTL 15
|||||

DB 1 PPTCSQANSGRISTL 15
|||||

RESULT 4
US-10-106-698-4719

Sequence 4719, Application US/10106698

Publication No. US20030109690A1

GENERAL INFORMATION:

APPLICANT: Ruben et al.

TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide

FILE REFERENCE: PA005P1

CURRENT APPLICATION NUMBER: US/10/106,698

PRIOR FILING DATE: 2002-03-27

PRIOR APPLICATION NUMBER: PCT/US00/26524

PRIOR FILING DATE: 2000-09-28

PRIOR APPLICATION NUMBER: US 60/157,137

PRIOR FILING DATE: 1999-09-29

PRIOR APPLICATION NUMBER: US 60/163,280

PRIOR FILING DATE: 1999-11-03

NUMBER OF SEQ ID NOS: 8564

SOFTWARE: PatentIn Ver. 3.0

SEQ ID NO 4719

LENGTH: 77

TYPE: PRT

ORGANISM: Homo sapiens

US-10-106-698-4719

Query Match 51.9%; Score 41; DB 15; Length 77;
Best Local Similarity 63.6%; Pred. No. 32;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PPTCSQANSGRISTL 12
|||||

DB 59 PPTCSQANSGRISTL 69

RESULT 5
US-09-764-868-1068

Sequence 1068, Application US/09764868

Patent No. US20020168711A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PT232

CURRENT APPLICATION NUMBER: US/09/764,868

PRIOR FILING DATE: 2001-01-17

Prior application data removed - refer to PAM or file wrapper

NUMBER OF SEQ ID NOS: 1510

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1068

LENGTH: 133

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (130)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (46)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (67)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-764-868-1068

Query Match 51.9%; Score 41; DB 10; Length 133;
Best Local Similarity 77.8%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPTCSQANS 9
|||||

DB 9 PPTCSQANS 17

RESULT 6
US-09-955-999-91
; Sequence 91, Application US/09955999
; Publication No. US20030036505A1
; GENERAL INFORMATION:
; APPLICANT: Barash et al.
; TITLE OF INVENTION: Signal Transduction Pathway Component Polynucleotides, Polypeptide
; TITLE OF INVENTION: Antibodies, and Methods Based Thereon
; FILE REFERENCE: PTO86P1
; CURRENT APPLICATION NUMBER: US/09/955,999
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-955-999-91

Query Match 51.9%; Score 41; DB 11; Length 133;
Best Local Similarity 77.8%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPTCSQANS 9
Db 9 PPTCSQANS 17

RESULT 7
US-10-106-698-6252
; Sequence 6252, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6252
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (126)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6252

Query Match 50.6%; Score 40; DB 15; Length 126;
Best Local Similarity 50.0%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PPTCSQANSGR1 12
Db 11 PPTCSQANSGRV 22

RESULT 8
US-09-925-302-768
; Sequence 768, Application US/09925302
; Patent No. US2002004941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 768
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-768

Query Match 50.6%; Score 40; DB 9; Length 404;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PPTCSQANSGR1 12
Db 11 PPTCSQANSGRV 22

RESULT 9
US-09-925-300-1118
; Sequence 1118, Application US/09925300
; Patent No. US2002015161A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1118
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (45)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (80)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1118

Query Match 49.4%; Score 39; DB 10; Length 80;
Best Local Similarity 58.3%; Pred. No. 69;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PPTCSQANSGR1 12
Db 18 PPTLQAAAAGRL 29

RESULT 10
US-10-156-761-11117
; Sequence 11117, Application US/10156761.
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11117
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11117

Query Match 49.4%; Score 39; DB 15; Length 161;
Best Local Similarity 57.1%; Pred. No. 1.4e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 PPTCSQANSGR1 14
Db 52 PPTVCALSGQLST 65

RESULT 11
US-09-888-615-114
; Sequence 114, Application US/09888615
; Patent No. US2002064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARVDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 114
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-114

Query Match 49.4%; Score 39; DB 9; Length 359;
Best Local Similarity 46.7%; Pred. No. 3.1e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 PPTCSQANSGR1 15

Db 3 PPSGSPASKMRIGLL 17

RESULT 12
US-09-771-161A-145
; Sequence 145, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 145
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-145

Query Match 49.4%; Score 39; DB 10; Length 401;
Best Local Similarity 50.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 PPTCSQANSGR1 15
Db 118 PPTSATNLSRVAGL 131

RESULT 13
US-09-771-161A-144
; Sequence 144, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 144
; LENGTH: 780
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-144

Query Match 49.4%; Score 39; DB 10; Length 780;
Best Local Similarity 50.0%; Pred. No. 6.6e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 PPTCSQANSGR1 15
Db 118 PPTSATNLSRVAGL 131

RESULT 14
US-09-771-161A-235
; Sequence 235, Application US/09771161A

```

/ Patent No. US2002011081A1
/ GENERAL INFORMATION:
/ APPLICANT: LEVINE, et al.
/ TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
/ FILE REFERENCE: 802620-2005.1
/ CURRENT APPLICATION NUMBER: US/09/771,161A
/ CURRENT FILING DATE: 2001-01-26
/ PRIOR APPLICATION NUMBER: 09/724,676
/ PRIOR FILING DATE: 2000-11-28
/ PRIOR APPLICATION NUMBER: 136776
/ PRIOR FILING DATE: 2000-06-15
/ PRIOR APPLICATION NUMBER: 135619
/ PRIOR FILING DATE: 2000-04-12
/ NUMBER OF SEQ ID NOS: 273
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 235
/ LENGTH: 942
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-771-161A-235

```

```

Query Match          49.4%; Score 39; DB 10; Length 942;
Best Local Similarity 50.0%; Pred. No. 8e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      2 PTCSONSGRISTL 15
      |||||  ||:|
Db      118 PTCSATNLSRVAGL 131

```

```

RESULT 15
US-09-771-161A-236
/ Sequence 236, Application US/09771161A
/ Patent No. US2002011081A1
/ GENERAL INFORMATION:
/ APPLICANT: LEVINE, et al.
/ TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
/ FILE REFERENCE: 802620-2005.1
/ CURRENT APPLICATION NUMBER: US/09/771,161A
/ CURRENT FILING DATE: 2001-01-26
/ PRIOR APPLICATION NUMBER: 09/724,676
/ PRIOR FILING DATE: 2000-11-28
/ PRIOR APPLICATION NUMBER: 136776
/ PRIOR FILING DATE: 2000-06-15
/ PRIOR APPLICATION NUMBER: 135619
/ PRIOR FILING DATE: 2000-04-12
/ NUMBER OF SEQ ID NOS: 273
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 236
/ LENGTH: 942
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-771-161A-236

```

```

Query Match          49.4%; Score 39; DB 10; Length 942;
Best Local Similarity 50.0%; Pred. No. 8e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      2 PTCSONSGRISTL 15
      |||||  ||:|
Db      118 PTCSATNLSRVAGL 131

```

Search completed: November 26, 2003, 12:38:39
Job time : 27.1084 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:21:29 ; Search time 41.5663 Seconds
(Without alignments)
57.280 Million cell updates/sec

Title: US-09-230-111C-14

Perfect score: 79

Sequence: 1 PPTCGANGRISTL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	15	AAW50175	Signal-transducing
2	56	70.9	31	AAU86416	Androgen receptor
3	56	70.9	31	AAU86429	Androgen receptor
4	46	58.2	1082	ABG29167	Novel human diapo
5	44	55.7	87	AAU63049	Protonibacterium
6	44	55.7	409	AAW43449	Tobacco laccase cl
7	43	54.4	198	ABG02424	Novel human diapo
8	41	51.9	77	AAW73945	Human colon cancer
9	41	51.9	84	AAU44755	Protonibacterium

10	41	51.9	133	22	AAU17503	Novel signal trans
11	41	51.9	144	22	AAW39866	Human polypeptide
12	41	51.9	148	22	AAW41652	Human polypeptide
13	40	50.6	126	22	AAW75478	Human colon cancer
14	40	50.6	404	21	AAW58430	Lung cancer associ
15	40	50.6	404	22	AAW40624	Human polypeptide
16	40	50.6	557	19	AAW43448	Tobacco laccase cl
17	40	50.6	748	22	ABW60412	Drosophila melanog
18	39	49.4	72	22	AAU61304	Protonibacterium
19	39	49.4	80	21	AAW56540	Human prostate can
20	39	49.4	92	22	AAU41900	Protonibacterium
21	39	49.4	104	22	ABG04821	Novel human diapo
22	39	49.4	117	21	AAW70526	Matze plaacid tag
23	39	49.4	153	21	AAW71106	Human Hydrolyase pr
24	39	49.4	222	21	AAW44730	Zea mays protein E
25	39	49.4	359	23	AAU82756	Amino acid sequenc
26	39	49.4	542	22	AAW61779	Sunflower berberin
27	39	49.4	542	22	AAW61791	Novel human diapo
28	39	49.4	939	22	ABG12508	Novel human diapo
29	39	49.4	939	22	ABG14439	Novel human diapo
30	39	49.4	940	22	ABG07054	Novel human diapo
31	39	49.4	941	21	AAW94737	Protein kinase C r
32	39	49.4	942	19	AAW56699	Modified protein K
33	39	49.4	1074	22	ABG07890	Novel human diapo
34	39	49.4	1074	22	ABG10255	Novel human diapo
35	39	49.4	1257	20	AAW06427	Mouse circadian re
36	38.5	48.7	108	21	AAW40463	Human ORFX ORF27
37	38.5	48.7	108	23	ABP00490	Human ORFX protein
38	38.5	48.7	239	22	ABG27685	Novel human diapo
39	38	48.1	78	22	ABG27485	Novel human diapo
40	38	48.1	92	20	AAW98996	Urokinase-type pla
41	38	48.1	92	22	AAU04456	Human u-PAR extrac
42	38	48.1	92	22	AAW84735	Human immune/haema
43	38	48.1	94	22	AAW74950	Human colon cancer
44	38	48.1	123	17	AAW94712	PRRSV ISU-1894 nuc
45	38	48.1	123	17	AAW94713	PRRSV ISU-22 nucle

ALIGNMENTS

RESULT 1

AAW50175

ID AAW50175 standard; peptide; 15 AA.

XX

AC AAW50175;

XX

DT 16-JUL-1998 (first entry)

XX

DE Signal-transducing protein carboxy-terminal peptide.

XX

KX Inhibition; specific binding; signal-transducing protein;

KW cytoplasmic protein; proliferation; cancer cell; apoptosis;

KM virally infected cell.

XX

OS Synthetic.

XX

PN WO9805347-A1.

XX

PD 12-FEB-1998.

XX

PF 18-JUL-1997; 97WO-US12677.

XX

PR 22-JUL-1996; 96US-0681219.

XX

PA (UYCO) UNIV COLUMBIA, NEW YORK.

XX

PI Sato T, Yanagisawa J.

XX

DR WPI; 1998-145347/13.

XX

PT Inhibition of signal transduction - by inhibiting binding between a
signal-transducing protein and a cytoplasmic protein, for treating

PT e.g. cancer or viral infection
 XX
 PS Claim 17; Page 64; 108pp; English.
 XX
 CC A novel composition is capable of inhibiting specific binding
 CC between a signal-transducing protein (STP) having the
 CC carboxy-terminal sequence (Ser/Thr)-Xaa-(Val/Ile/Leu), where Xaa =
 CC any amino acid (e.g. the present peptide), and a cytoplasmic
 CC protein (CP) containing the sequence AAM50162 or AAM50163.
 CC The composition can be used to inhibit the proliferation of cancer
 CC or virally infected cells, or induce apoptosis in cancer or virally
 CC infected cells.
 CC
 SQ Sequence 15 AA;
 Query Match 100.0%; Score 79; DB 19; Length 15;
 Best Local Similarity 100.0%; Pred. No. 7.8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPTCSQANSGRISTL 15
 Db 1 PPTCSQANSGRISTL 15
 RESULT 2
 ID AAM50162 standard; Peptide; 31 AA.
 AC AAM50162;
 XX
 XX 21-MAY-2002 (first entry)
 DT
 DE Androgen receptor binding peptide B8H3.
 XX
 XX Androgen receptor; breast cancer; combinatorial peptide library;
 KW receptor modulating compound.
 XX
 OS Synthetic.
 XX
 PN WO200204956-A2.
 XX
 PD 17-JAN-2002.
 XX
 PF 11-JUL-2001; 2001WO-US21867.
 XX
 PR 12-JUL-2000; 2000US-0614865.
 XX
 PR 21-MAY-2001; 2001US-0860688.
 XX
 PA (KARO-) KARO BIO USA INC.
 XX
 PI Fowlkes DM, Barnett TR, Buehrer B;
 XX
 DR WPI; 2002-154969/20.
 XX
 PT Identifying receptor-binding peptides comprises screening combinatorial
 PT peptide library presented in form of cells each of which coexpress one
 PT peptide member and receptor with signal producing system for reporting
 PT binding -
 XX
 XX Example 502; Page 154; 175pp; English.
 PS
 XX The invention relates to identifying a binding peptide which binds a
 CC receptor and which is a member of a combinatorial library of peptides,
 CC comprising screening a combinatorial peptide library presented in the
 CC form of cells which coexpress the receptor or its ligand-binding receptor
 CC moiety and one member of the library, together with a signal producing
 CC system for reporting binding of the peptide to the receptor. Also
 CC included is a method for predicting the receptor-modulating activity of a
 CC compound which modulates the biological activity of a receptor
 CC comprising (a) identifying peptides which bind the receptor by the
 CC method above, (b) using a number of the peptides to predict the receptor-
 CC modulating activity of a compound by (i) providing a panel of
 CC identified peptides, where the members differ in their ability to bind

CC to the receptor depending on reference conformations the receptor is
 CC in, where the effect of a number of reference substances known to
 CC modulate the biological activity of the receptor on the binding of each
 CC member of the panel is known and is characterised as a reference
 CC fingerprint for each reference substance, (ii) screening a test substance
 CC of unknown activity relative to the receptor to determine its effect on
 CC the binding of each member of the panel to the receptor, thereby
 CC obtaining a test fingerprint for the test substance, (iii) comparing the
 CC test fingerprint to the reference fingerprints and (iv) predicting the
 CC biological activity of the test substance based on the assumption that
 CC its biological activity will be similar to that of reference substances
 CC with similar fingerprints. The method is useful for identifying a binding
 CC peptide which binds a vertebrate, mammalian, preferably human receptor,
 CC an intracellular, nuclear, oestrogen or androgen receptor. The identified
 CC peptides which bind to the receptor are useful for predicting the
 CC receptor-modulating activity of a compound (e.g. ant/agonists).
 CC The receptor-binding library members are useful in the prediction of the
 CC ability of small organic molecules, suitable for pharmaceutical use
 CC (e.g. in the case of oestrogen receptors, for breast cancer treatment),
 CC to interact with the receptor. The analyte-binding molecules can also be
 CC used for in vivo imaging. The method has several advantages over whole
 CC animal-based assay systems in that the same technology can be applied to
 CC a variety of different receptors, the system can be used for high
 CC throughput screening and compound characterisation, and gives very
 CC distinct patterns for agonists and antagonists of receptor activity using
 CC very much less protein. The present sequence is an androgen receptor
 CC binding peptide from a combinatorial peptide library.
 CC
 SQ Sequence 31 AA;
 Query Match 70.9%; Score 56; DB 23; Length 31;
 Best Local Similarity 100.0%; Pred. No. 0.074;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 TCSQANSGRIS 13
 Db 18 TCSQANSGRIS 28
 RESULT 3
 ID AAM50162 standard; Peptide; 31 AA.
 AC AAM50162;
 XX
 XX 21-MAY-2002 (first entry)
 DT
 DE Androgen receptor binding peptide B8H3.
 XX
 XX Androgen receptor; breast cancer; combinatorial peptide library;
 KW receptor modulating compound.
 XX
 OS Synthetic.
 XX
 PN WO200204956-A2.
 XX
 PD 17-JAN-2002.
 XX
 PF 11-JUL-2001; 2001WO-US21867.
 XX
 PR 12-JUL-2000; 2000US-0614865.
 XX
 PR 21-MAY-2001; 2001US-0860688.
 XX
 PA (KARO-) KARO BIO USA INC.
 XX
 PI Fowlkes DM, Barnett TR, Buehrer B;
 XX
 DR WPI; 2002-154969/20.
 XX
 PT Identifying receptor-binding peptides comprises screening combinatorial
 PT peptide library presented in form of cells each of which coexpress one
 PT peptide member and receptor with signal producing system for reporting
 PT binding -

XX Example 502; Page 155; 175pp; English.

CC The invention relates to identifying a binding peptide which binds a
 CC receptor and which is a member of a combinatorial library of peptides,
 CC comprising screening a combinatorial peptide library presented in the
 CC form of cells which coexpress the receptor or its ligand-binding receptor
 CC moiety and one member of the library, together with a signal producing
 CC system for reporting binding of the peptide to the receptor. Also
 CC included is a method for predicting the receptor-modulating activity of a
 CC compound which modulates the biological activity of a receptor
 CC comprising (a) identifying peptides which bind the receptor by the
 CC method above, (b) using a number of the peptides to predict the receptor-
 CC modulating activity of a compound by (i) providing a panel of
 CC identified peptides, where the members differ in their ability to bind
 CC to the receptor depending on reference conformations the receptor is
 CC in, where the effect of a number of reference substances known to
 CC modulate the biological activity of the receptor on the binding of each
 CC member of the panel is known and is characterized as a reference
 CC fingerprint for each reference substance, (ii) screening a test substance
 CC of unknown activity relative to the receptor to determine its effect on
 CC the binding of each member of the panel to the receptor, thereby
 CC obtaining a test fingerprint for the test substance, (iii) comparing the
 CC test fingerprint to the reference fingerprints and (iv) predicting the
 CC biological activity of the test substance based on the assumption that
 CC its biological activity will be similar to that of reference substances
 CC with similar fingerprints. The method is useful for identifying a binding
 CC peptide which binds a vertebrate, mammalian, preferably human receptor,
 CC an intracellular, nuclear, oestrogen or androgen receptor. The identified
 CC peptides which bind to the receptor are useful for predicting the
 CC receptor-modulating activity of a compound (e.g. antiagonists).
 CC The receptor-binding library members are useful in the prediction of the
 CC ability of small organic molecules, suitable for pharmaceutical use
 CC (e.g. in the case of oestrogen receptors, for breast cancer treatment),
 CC to interact with the receptor. The analyte-binding molecules can also be
 CC used for in vivo imaging. The method has several advantages over whole
 CC animal-based assay systems in that the same technology can be applied to
 CC a variety of different receptors, the system can be used for high
 CC throughput screening and compound characterisation, and gives very
 CC distinct patterns for agonists and antagonists of receptor activity using
 CC very much less protein. The present sequence is an androgen receptor
 CC binding peptide from a combinatorial peptide library.

SO Sequence 31 AA;

Query Match 70.9%; Score 56; DB 23; Length 31;
 Best Local Similarity 100.0%; Pred. No. 0.074;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TCSQANSGRIS 13
 |||||
 Db 18 TCSQANSGRIS 28

RESULT 4

ABG29167 standard; Protein; 1082 AA.

AC ABG29167;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #29158.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Dymenac RT, Liu C, Tang YT;

DR WPI, 2001-639362/73.

DR N-PSDB; AAS93354.

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT biologically for genetic disorders or other traits and to assess

PS Claim 20; SEQ ID No 59526; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II), and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1082 AA;

Query Match 58.2%; Score 46; DB 22; Length 1082;
 Best Local Similarity 64.3%; Pred. No. 99;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 PPTCSQANSGRIST 14
 |||||
 Db 379 PPTCSQANSGRIST 392

RESULT 5

AAU63049 standard; Protein; 87 AA.

AC AAU63049;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #23945.

KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;

KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

OS Propionibacterium acnes.

XX WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59631.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 PS Example 1; SEQ ID No 24244; 1069pp; English.
 XX
 CC Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 87 AA;
 XX
 Query Match 55.7%; Score 44; DB 22; Length 87;
 Best Local Similarity 42.9%; Pred. No. 17;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
 QY 1 PPTCSQANSGRIST 14
 Db 25 PPTCSQANSSEVNS 38
 XX
 RESULT 6
 AAW43449
 ID AAW43449 standard; Protein; 409 AA.
 XX
 AC AAW43449;
 XX
 DT 01-JUN-1998 (first entry)
 XX
 DE Tobacco laccase clone pTL-2 protein.
 XX
 KW Tobacco; laccase; oxidoreductase; probe; hybridisation; sycamore;
 KW transgenic plant; digestibility; forage crop; paper manufacture;
 KW lignin; parasite resistance.
 XX
 OS Nicotiana tabacum.
 XX
 PN WO9745549-A1.
 XX
 PD 04-DEC-1997.
 XX
 PF 30-MAY-1997; 97WO-FR00948.

XX
 PR 31-MAY-1996; 96FR-0006760.
 XX
 PA (CNRS) CENT NAT RECH SCT.
 XX
 PI Faye L, Gomord VM, Kiefer-Meyer MC, O'Connell A;
 DR WPI; 1998-032655/03.
 DR N-PSDB; AA01597.
 XX
 PT Modifying lignin biosynthesis in plants with gene encoding laccase
 PT mRNA - or its anti-sense complement, especially for crops used as
 PT fodder and for paper production
 XX
 PS Claim 19; Page 37-39; 72pp; French.
 XX
 CC This sequence represents a tobacco laccase (an oxidoreductase acting on
 CC diphenols) encoded by the partial length clone pTL-2. The sequence
 CC is lacking part of the C-terminus of the protein. The gene sequence was
 CC isolated from a cDNA library from tobacco leaves using a probe derived
 CC from the sequence of the sycamore (Acer pseudoplatanus) laccase gene.
 CC The gene can be used to transform plant cells for producing transgenic
 CC plants having a reduced or altered lignin content, for improved
 CC digestibility of forage crops, for paper manufacture. Plants with
 CC increased lignin content may have better resistance to parasites.
 XX
 SQ Sequence 409 AA;
 XX
 Query Match 55.7%; Score 44; DB 19; Length 409;
 Best Local Similarity 63.6%; Pred. No. 79;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 PPTCSQANSGRIST 12
 Db 218 PPTCSQANSRV 228
 XX
 RESULT 7
 ABG02424
 ID ABG02424 standard; Protein; 198 AA.
 XX
 AC ABG02424;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #2415.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS66611.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX

PS Claim 20; SEQ ID No 32783; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG3037 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX
 XX SQ Sequence 198 AA;

Query Match 54.4%; Score 43; DB 22; Length 198;
 Best Local Similarity 52.6%; Pred. No. 55;
 Matches 10; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

QY 1 PPTCSQ----ANSGRISTL 15
 |||||
 DB 143 PPLCGSGPARAGSRINAL 161

RESULT 8
 AAG73945
 ID AAG73945 standard; Protein; 77 AA.

XX AAG73945;
 XX
 XX DT 03-SEP-2001 (first entry)
 XX
 XX DE Human colon cancer antigen protein SEQ ID NO:4709.
 XX
 XX KM Human; colon cancer; colon cancer antigen; diagnosis; detection;
 XX colorectal carcinoma.
 XX
 XX OS Homo sapiens.
 XX
 XX PN WO200122920-A2.
 XX
 XX PD 05-APR-2001.
 XX
 XX PF 28-SEP-2000; 2000WO-US26524.
 XX
 XX PR 29-SEP-1999; 99US-0157137.
 XX
 XX PR 03-NOV-1999; 99US-0163280.
 XX
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX
 XX DR WPI; 2001-235357/24.
 XX
 XX DR N-PSDB; AAH33376.
 XX
 XX PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 XX useful for preventing, diagnosing and/or treating colorectal cancers -
 XX
 XX PS Claim 11; Page 6511; 9803pp; English.
 XX
 XX CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 XX cancer-associated nucleic acid molecules (N) and proteins (P), where

CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated P,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77269 represent sequences used in the exemplification of the
 CC present invention.

CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX
 XX SQ Sequence 77 AA;

Query Match 51.9%; Score 41; DB 22; Length 77;
 Best Local Similarity 63.6%; Pred. No. 45;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PTCQANSGRI 12
 |||||
 DB 59 PACSQAGSGAV 69

RESULT 9
 AAU44755
 ID AAU44755 standard; Protein; 84 AA.

XX AAU44755;
 XX
 XX DT 27-FEB-2002 (first entry)
 XX
 XX DE Propionibacterium acnes immunogenic protein #5651.
 XX
 XX KM SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 XX uveitis; endophthalmitis; bone; joint; central nervous system; ERISA;
 XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 XX dermatological; osteopathic; neuroprotectant.
 XX
 XX OS Propionibacterium acnes.
 XX
 XX PN WO200181581-A2.
 XX
 XX PD 01-NOV-2001.
 XX
 XX PF 20-APR-2001; 2001WO-US12865.
 XX
 XX PR 21-APR-2000; 2000US-199047P.
 XX
 XX PR 02-JUN-2000; 2000US-208841P.
 XX
 XX PR 07-JUL-2000; 2000US-216747P.
 XX
 XX PA (COR-) CORIXA CORP.
 XX
 XX PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhacia A;
 XX
 XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 XX DR WPI; 2001-616774/71.
 XX
 XX DR N-PSDB; AAS59524.
 XX
 XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
 XX vaccinating against and diagnosing infections, especially useful for
 XX treating acne vulgaris -
 XX
 XX PS Example 1; SEQ ID No 5950; 1069pp; English.
 XX
 XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 XX polypeptides. The proteins and their associated DNA sequences are used in
 XX the treatment, prevention and diagnosis of medical conditions caused by

PR	14-JUL-2000	2000US-0218299
PR	26-JUL-2000	2000US-0220363
PR	26-JUL-2000	2000US-0220364
PR	14-AUG-2000	2000US-0224518
PR	14-AUG-2000	2000US-0224519
PR	14-AUG-2000	2000US-0225213
PR	14-AUG-2000	2000US-0225214
PR	14-AUG-2000	2000US-0225266
PR	14-AUG-2000	2000US-0225267
PR	14-AUG-2000	2000US-0225268
PR	14-AUG-2000	2000US-0225270
PR	14-AUG-2000	2000US-0225447
PR	14-AUG-2000	2000US-0225757
PR	14-AUG-2000	2000US-0225758
PR	14-AUG-2000	2000US-0225759
PR	18-AUG-2000	2000US-0226279
PR	22-AUG-2000	2000US-0226661
PR	22-AUG-2000	2000US-0226668
PR	22-AUG-2000	2000US-0227182
PR	23-AUG-2000	2000US-0227009
PR	30-AUG-2000	2000US-0228924
PR	01-SEP-2000	2000US-0229287
PR	01-SEP-2000	2000US-0229343
PR	01-SEP-2000	2000US-0229344
PR	01-SEP-2000	2000US-0229345
PR	03-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0231342
PR	08-SEP-2000	2000US-0231343
PR	08-SEP-2000	2000US-0231444
PR	08-SEP-2000	2000US-0231445
PR	08-SEP-2000	2000US-0231413
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0232080
PR	08-SEP-2000	2000US-0232081
PR	12-SEP-2000	2000US-0231368
PR	14-SEP-2000	2000US-0232397
PR	14-SEP-2000	2000US-0232398
PR	14-SEP-2000	2000US-0232399
PR	14-SEP-2000	2000US-0232400
PR	14-SEP-2000	2000US-0232401
PR	14-SEP-2000	2000US-0233063
PR	14-SEP-2000	2000US-0233064
PR	14-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0234223
PR	21-SEP-2000	2000US-0234274
PR	25-SEP-2000	2000US-0234998
PR	25-SEP-2000	2000US-0234998
PR	26-SEP-2000	2000US-0235484
PR	27-SEP-2000	2000US-0235834
PR	27-SEP-2000	2000US-0235836
PR	29-SEP-2000	2000US-0236327
PR	29-SEP-2000	2000US-0236367
PR	29-SEP-2000	2000US-0236368
PR	29-SEP-2000	2000US-0236670
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237039
PR	02-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239353
PR	13-OCT-2000	2000US-0239397
PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241221
PR	20-OCT-2000	2000US-0241785
PR	20-OCT-2000	2000US-0241786
PR	20-OCT-2000	2000US-0241787
PR	20-OCT-2000	2000US-0241809
PR	20-OCT-2000	2000US-0241809
PR	01-NOV-2000	2000US-0244517

PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251038.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251889.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465460/50.
XX
XX N-PSDB; AAS67420.
XX
XX Novel polypeptides useful for diagnosing, treating, preventing and/or
XX
XX PT prognosing disorders related to the proteins, including cancers, immune
XX
XX disorders and neuronal disorders -
XX
XX
XX Claim 1; SEQ ID NO 1068; 880bp; English.

CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorder), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
CC pathway protein, amino acid sequences of the invention.
XX
Query Match 51.9%; Score 41; DB 22; Length 133;
Best Local Similarity 77.8%; Pred. No. 77;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 PPTCSQANS 9
||:|||||
DB 9 PPTCSQANS 17
RESULT 11
AAM39866
ID AAM39866 standard; Protein; 144 AA.
XX
XX AAM39866;
AC
XX
XX
DT 22-OCT-2001 (first entry)
XX
XX
DE Human polypeptide SEQ ID NO 3011.
XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheric nerve system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.
XX
XX Homo sapiens.
OS
XX
XX MO200153312-A1.
PN
XX
XX
PD 26-JUL-2001.
XX
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX
XX 21-JAN-2000; 2000US-0488725.
XX
XX 25-APR-2000; 2000US-0552317.
XX
XX 09-JUL-2000; 2000US-0598042.
XX
XX 19-JUL-2000; 2000US-0620312.
XX
XX 03-AUG-2000; 2000US-0653450.
XX
XX 14-SEP-2000; 2000US-0662191.
XX
XX 19-OCT-2000; 2000US-0693036.
XX
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
XX
XX N-PSDB; AA159022.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX
XX PT such as central nervous system injuries -
XX
XX
XX Example 4; SEQ ID NO 3011; 10078bp; English.
XX
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX
XX CC the encoded polypeptides (AAM36642-AA42213) with nootropic,
XX
XX immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activation/inhibit activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX SQ Sequence 144 AA;

Query Match 51.9%; Score 41; DB 22; Length 144;

Best Local Similarity 63.6%; Pred. No. 84;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PTCSQANSGR1 12

Db 126 PACSQAGSGAV 136.

RESULT 12

ID AAM41652 standard; Protein; 148 AA.

XX AAM41652;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6583.

XX Human; nocertopic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

XX 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 03-AUG-2000; 2000US-0620312.

PR 14-SEP-2000; 2000US-0653450.

PR 19-OCT-2000; 2000US-0662191.

PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Ahmadi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX N-PSDB; AA160808.

XX Novel nucleic acids and polypeptides, useful for treating disorders

XX such as central nervous system injuries -

XX Example 2; SEQ ID NO 6583; 10078bp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAG75478-AA42213) with nocertopic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activation/inhibit activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX SQ Sequence 148 AA;

Query Match 51.9%; Score 41; DB 22; Length 148;

Best Local Similarity 63.6%; Pred. No. 86;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PTCSQANSGR1 12

Db 130 PACSQAGSGAV 140

RESULT 13

ID AAG75478 standard; Protein; 126 AA.

XX AAG75478;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen protein SEQ ID NO:6242.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

KW colorectal carcinoma.

OS Homo sapiens.

PN WO200122920-A2.

XX 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX N-PSDB; AA134883.

XX Claim 11; Page 7694; 9803pp; English.

XX AAH32943 to AAH37195 and AAG77788 represent human colon

XX cancer-associated nucleic acid molecules (N) and proteins (P), where

XX the proteins are collectively known as colon cancer antigens. The colon

XX cancer antigens have cytostatic activity and can be used in gene

XX therapy and vaccine production. N and P may be used in the prevention,

XX diagnosis and treatment of diseases associated with inappropriate P

XX expression. For example, N and P may be used to treat disorders

XX associated with decreased expression by rectifying mutations or deletions

XX in a patient's genome that affect the activity of P by expressing

CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAH7789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX
SQ Sequence 126 AA;

Query Match 50.6%; Score 40; DB 22; Length 126;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 PPTCSQANSGR1 12
Db 11 PPRCEAASQGRV 22
||| |:
||| |:
||| |:

RESULT 14
AAB58430
ID AAB58430 standard; Protein; 404 AA.
XX
XX AAB58430,
XX
XX 14-MAR-2001 (first entry)
XX
XX Lung cancer associated polypeptide sequence SEQ ID 768.
XX
XX Human; lung cancer associated protein; neuroprotective; cytostatic;
XX cardioactive; immunomodulatory; muscular active; vulnery;
XX gastrointestinal; nephroretropic; antiinfective; gynecological;
XX antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
XX proliferative disorder; wound healing; infectious disease.
XX
XX Homo sapiens.
XX
XX WO200055180-A2.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000MO-US05918.
XX
XX 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX
XX Ruben SM;
XX
XX MPI; 2000-587514/55.
XX N-PSDB; AAF18306.
XX
XX Lung cancer associated gene sequences, referred to as lung cancer
XX antigens, useful for treatment, prevention, and diagnosis of disorders
XX such as lung cancer -
XX
XX Claim 11; Page 1292-1293; 1425pp; English.
XX
XX polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
XX associated proteins represented in AAB58106 - AAB58548. Lung cancer
XX associated proteins and polynucleotide sequences, their agonists, and
XX antagonists may have neuroprotective; cytostatic; cardioactive;
XX immunomodulatory; muscular active general; vulnery; gastrointestinal
XX general; nephroretropic; antiinfective; gynecological; or antibacterial
XX activity. The invention also includes antibodies specific for the
XX protein or polynucleotide sequences. The lung cancer associated
XX polynucleotide sequences may be used for detection of lung cancer,
XX chromosome identification, as chromosome markers, and for numerous other
XX diagnostic or research purposes. The proteins may be used to treat

CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders. The proteins may also be used in the treatment of wounds and
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
CC peptide AAB58549 are used in the course of the invention for the
CC identification and characterization of the polynucleotide and protein
CC sequences.
XX
XX
SQ Sequence 404 AA;

Query Match 50.6%; Score 40; DB 21; Length 404;
Best Local Similarity 50.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 PPTCSQANSGR1 12
Db 11 PPRCEAASQGRV 22
||| |:
||| |:
||| |:

RESULT 15
AAM40624
ID AAM40624 standard; Protein; 404 AA.
XX
XX AAM40624;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 5555.
XX
XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX leukemetic; thrombolytic; drug screening; arthritis; inflammation;
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000MO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao Q, Zhou P, Goodrich R, Dirmanc RT;
XX
XX MPI; 2001-442253/47.
XX N-PSDB; AAI59780.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 5555; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX the encoded polypeptides (AAM38642-AA44213) with noctropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX

SQ Sequence 404 AA;

Query Match 50.6%; Score 40; DB 22; Length 404;

Best Local Similarity 50.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PPTCSQANSGR1 12

Db 11 PPRCEAASQGRV 22

Search completed: November 26, 2003, 12:30:25
Job time : 42.5663 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 26, 2003, 12:22:05 ; Search time 7.22892 Seconds
(without alignment)
97.580 Million cell updates/sec

Title: US-09-230-111C-14

Perfect score: 79
Sequence: 1 PPTCSQANSGRISTL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	50.6	474	1	MIG1_KLULA
2	39	49.4	109	1	P20898 Kluveromyc
3	39	49.4	942	1	P02539 ovine aries
4	39	49.4	951	1	PKL1_HUMAN
5	39	49.4	951	1	SPR8_HUMAN
6	38	48.1	272	1	PER1_MOUSE
7	38	48.1	272	1	NIH2_CLOPA
8	38	48.1	335	1	NIH6_CLOPA
9	38	48.1	348	1	UPAR_HUMAN
10	38	48.1	428	1	Y479_WYCTU
11	38	48.1	428	1	MTT8_THETH
12	38	48.1	474	1	ANA_DROME
13	38	48.1	584	1	COE1_RAT
14	38	48.1	586	1	COE1_HUMAN
15	38	48.1	591	1	COE1_MOUSE
16	38	48.1	1966	1	CCAF_HUMAN
17	38	48.1	3375	1	UN52_CAEEL
18	37.5	47.5	349	1	LACH_SCHAM
19	37	46.8	447	1	OPAG_NEIGO
20	37	46.8	288	1	T2G1_HAEGA
21	37	46.8	492	1	TISD_HUMAN
22	37	46.8	564	1	AGGL_RICCO
23	37	46.8	596	1	COE3_HUMAN
24	37	46.8	596	1	COE3_MOUSE
25	37	46.8	703	1	NHR3_CAEEL
26	37	46.8	1852	1	CCAS_CYPCA
27	36	45.6	93	1	YEHE_ECOLI
28	36	45.6	295	1	NIH1_RHOCA
29	36	45.6	337	1	GOOS_YEAST
30	36	45.6	339	1	LPXD_CAUCR
31	36	45.6	490	1	FXN3_HUMAN
32	36	45.6	491	1	K2M2_SHEEP
33	36	45.6	548	1	ERF_HUMAN
			551	1	ERF_MOUSE

ALIGNMENTS

34	36	45.6	576	1	RIC1_RICCO	P02879 ricinus com
35	36	45.6	598	1	COE3_XENLA	O73742 xenopus lae
36	36	45.6	1080	1	SET1_YEAST	P38827 saccharomyc
37	36	45.6	1140	1	YD76_MYCPN	P75405 mycoplasma
38	36	45.6	3726	1	TRX_DROME	P20659 drosophilla
39	35	44.3	108	1	KAC_RANCA	P11272 rana catesb
40	35	44.3	301	1	YAB1_SCHPO	O09806 schizosacch
41	35	44.3	379	1	ROM3_HUMAN	P51991 homo sapien
42	35	44.3	396	1	FLGE_BRUAB	O52070 brucei ab
43	35	44.3	407	1	FMDA_METME	O50228 methyllophi
44	35	44.3	419	1	KCRS_RABIT	O77814 oryctolagus
45	35	44.3	423	1	LE21_PYRAB	O9u07 pyrococcus

RESULT 1

ID MIG1_KLULA STANDARD; PRT; 474 AA.

AC P50898; 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Regulatory protein MIG1.

GN MIG1.

OS Kluveromyces fragilis (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OX Saccharomycetales; Saccharomycetaceae; Kluveromyces.

NCBI_Taxid=28985;

RE SEQUENCE FROM N.A.

RC STRAIN=CBS 2359 / IFO 1267 / NRRL Y-1140;

RA MEDLINE=95402208; PubMed=7672126;

RT "The MIG1 repressor from Kluveromyces fragilis: cloning, sequencing

and functional analysis in Saccharomyces cerevisiae."

RL FEMS Lett. 371:191-194 (1995).

CC -1- FUNCTION: INVOLVED IN GLUCOSE REPRESSION OF GLUCOSE METABOLISM

CC GENES.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- SIMILARITY: BELONGS TO THE CREM/MIG GROUP OF C2H2-TYPE ZINC-

CC FINGERS PROTEINS.

CC -1- SIMILARITY: Contains 2 C2H2-type zinc fingers.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL; 250017; CAA90320.1; -

DR PIR; S66480; S66480.

DR HSSP; P08047; ISE2.

DR InterPro; IPR007087; Znf_C2H2.

DR Pfam; PF000096; zfc_C2H2; 2.

DR ProDom; PD000003; Znf_C2H2; 2.

DR SMART; SM00355; ZNF_C2H2; 2.

DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.

DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 2.

KW Transcription regulation; Repressor; DNA-binding; Nuclear protein;

FT Zinc-finger; Metal-binding; Repeat; Carbohydrate metabolism.

FT ZN_FING 26 48 C2H2-TYPE 1.

FT ZN_FING 54 78 C2H2-TYPE 2.

FT DOMAIN 217 225 POLY-SER.

SQ SEQUENCE 474 AA; 52907 MW; 4A5F3EA3603B17C3 CRC64;

Query Match 50.6%; Score 40; DB 1; Length 474;

Best Local Similarity 60.0%; Pred. No. 19;

Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 PPTCSQANGSRISTL 15
 |||||
 DB 178 PPIRSGNNGNIDLL 192

RESULT 2
 K2M1 SHEEP STANDARD; PRT; 109 AA.
 ID P02539;
 AC 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 NCBI_TaxId=9940;
 [1]
 RP SEQUENCE.
 RX MEDLINE=79020757; PubMed=581264;
 RA Creweh W.G., Inglis A.S., McKern N.M.;
 RT "Amino acid sequences of alpha-helical segments from S-
 RL carboxymethylkeratine-A. Complete sequence of a type-II segment.";
 RL Biochem. J. 173:365-371(1978).
 CC -1- FUNCTION: WOOL MICROFIBRILLAR KERATIN.
 CC -1- MICELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
 CC MICROFIBRILLAR KERATIN, I AND II (40-55 AND 56-70 KILODALTONS,
 CC RESPECTIVELY). AT LEAST ONE TYPE OF CYTOSKELETAL KERATIN IS
 CC PRESENT IN ALL VERTEBRATE EPITHELIAL CELLS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC InterPro: IPR001664; IF.
 DR InterPro: IPR003054; Keratin_II.
 DR Pfam: PF00038; filament_1.
 DR PRINTS: PR01276; TYPE2KERATIN.
 DR PROSITE: PS00226; IF; PARTIAL.
 KW Intermediate filament; Coiled coil; Keratin.
 FT NON_TER 1 1
 FT DOMAIN <1 >109 ROD.
 FT LINKER 1.
 FT DOMAIN 11 >109 COIL 1B.
 FT VARIANT 51 51 R -> K.
 FT VARIANT 52 52 R -> K.
 FT NON_TER 109 109
 SQ SEQUENCE 109 AA; 12683 MW; 679DDDA5A621CF8 CRC64;

Query Match 49.4%; Score 39; DB 1; Length 109;
 Best Local Similarity 54.5%; Pred. No. 6.1;
 Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 4 CSQANGSRIST 14
 |||||
 DB 27 CABADSGRLSS 37

RESULT 3
 PKL1 HUMAN STANDARD; PRT; 942 AA.
 ID Q16512; Q15143;
 AC Q16512; Q15143;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DE 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Protein kinase C-like 1 (EC 2.7.1.-) (Protein-kinase C-related kinase
 DE 1) (Protein kinase C-like PKN) (Serine-threonine protein kinase N).
 GN PRKCL1 OR PRK1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxId=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=95080426; PubMed=7988719;
 RX Palmer R.H., Ridden J., Parker P.J.;

RT "Identification of multiple, novel, protein kinase C-related gene
 RT products.";
 RT FEBS Lett. 356:5-8(1994).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Fetal Brain;
 RA MEDLINE=95154310; PubMed=7851406;
 RX Palmer R.H., Ridden J., Parker P.J.;
 RT "Cloning and expression patterns of two members of a novel protein-
 RT kinase-C-related kinase family";
 RL Eur. J. Biochem. 227:344-351(1995).
 RN [3]
 RP SEQUENCE FROM N.A., AND MUTAGENESIS OF LYS-644.
 RC TISSUE=Hippocampus;
 RX MEDLINE=94183274; PubMed=8135837;
 RA Mukai H., Ono Y.;
 RT "A novel protein kinase with leucine zipper-like sequences: its
 RT catalytic domain is highly homologous to that of protein kinase C";
 RL Biochem. Biophys. Res. Commun. 199:897-904(1994).
 CC -1- FUNCTION: CAN PHOSPHORYLATE RIBOSOMAL PROTEIN S6. MEDIATES GTPASE
 CC RHO DEPENDENT INTRACELLULAR SIGNALING (BY SIMILARITY).
 CC -1- ENZYME REGULATION: ACTIVATED BY LIPIDS, PARTICULARLY CARDIOLIPIN
 CC AND TO A LESSER EXTENT BY OTHER ACIDIC PHOSPHOLIPIDS (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- TISSUE SPECIFICITY: FOUND UBICUOUSLY. EXPRESSED IN HEART, BRAIN,
 CC PLACENTA, LUNG, SKELETAL MUSCLE, KIDNEY, AND PANCREAS.
 CC -1- PTM: AUTOPHOSPHORYLATED. PREFERABLY IN SERINE.
 CC -1- PTM: ACTIVATED BY LIMITED PROTEOLYSIS WITH TRYPSIN (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC PKC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U33053; AAC50209.1; -
 CC EMBL; S75546; AAB33345.1; -
 CC EMBL; D26181; BAA05169.1; -
 CC PIR; JC2129; JC2129.
 CC PDB; 1CXZ; 23-MAY-00.
 CC GeneW; HGNC:9405; PRKCL1.
 CC MIM; 601032; -
 CC GO; GO:0007257; P:activation of JUN kinase; TAS.
 CC GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
 CC GO; GO:0007165; P:signal transduction; TAS.
 CC InterPro; IPR000008; C2.
 CC InterPro; IPR000961; Pkinase_C.
 CC InterPro; IPR000719; Prot_Kinase.
 CC InterPro; IPR000861; REM_repeat.
 CC InterPro; IPR002290; Ser_thr_kinase.
 CC Pfam; PF02185; HRI; 3.
 CC Pfam; PF00069; pkinase; 1.
 CC Pfam; PF00433; pkinase_C1.
 CC ProDom; PD000001; Prot_Kinase; 1.
 CC SMART; SM00239; C2; 1.
 CC SMART; SM00074; HRI; 3.
 CC SMART; SM00133; S_TK_X; 1.
 CC SMART; SM00220; S_TKC; 1.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 CC TRANSFERASE; ATP-binding; Serine/threonine-protein kinase;
 KW Phosphorylation; Polymorphism; 3d-structure.
 FT DOMAIN 615 874 PROTEIN KINASE.
 FT NP_BIND 621 629 ATP (BY SIMILARITY).
 FT BINDING 644 644 ATP (BY SIMILARITY).
 FT ACT_SITE 740 740 BY SIMILARITY.

FT VARIANT 901 901 V -> I (IN dbSNP:10846).
 FT MUTAGEN 644 644 K->R: SUBSTANTIAL REDUCTION OF
 FT CONFLICT 191 191 AUTOPHOSPHORYLATION.
 FT SEQUENCE 942 AA; 103989 MW; A89E40DCAEF560E3 CRC64;
 Query Match 49.4%; Score 39; DB 1; Length 942;
 Best Local Similarity 50.0%; Pred. No. 57;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 PTCQANSGRISTL 15
 Db 118 PTCSATNLSRVAGL 131

RESULT 4
 SFR8_HUMAN STANDARD; PRT; 951 AA.
 ID_SFR8_HUMAN 012872;
 AC 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Splicing factor, arginine/serine-rich 8 (Suppressor of white apricot
 protein homolog).
 GN SFRS8 OR SWAP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA MEDLINE=94266805; PubMed=8206918;
 RA Demhez F., Lafyatis R.;
 RT "Conservation of regulated alternative splicing and identification of
 functional domains in vertebrate homologs to the Drosophila splicing
 RT regulator suppressor-of-white-apricot";
 RL J. Biol. Chem. 269:16170-16179(1994).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=97094873; PubMed=8940107;
 RA Sarkisian M., Milne A., Lafyatis R.;
 RT "The mammalian homolog of suppressor-of-white-apricot regulates
 alternative mRNA splicing of CD45 exon 4 and fibronectin ITCS";
 RL J. Biol. Chem. 271:31106-31114(1996).
 CC -1- FUNCTION: MAY FUNCTION AS AN ALTERNATIVE SPLICING REGULATOR.
 CC REGULATE ITS OWN EXPRESSION AT THE LEVEL OF RNA PROCESSING. ALSO
 CC REGULATE THE SPLICING OF FIBRONECTIN AND CD45 GENES. MAY ACT, AT
 CC LEAST IN PART, BY INTERACTION WITH OTHER R/S-CONTAINING SPLICING
 CC FACTORS.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- SIMILARITY: Contains 2 SURP motif repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL, U08377; AAA19604.1; -
 DR Genew; HGNC:10790; SFRS8.
 DR GK; Q12872; -
 DR MIM; 601945; -
 DR GO; GO:0008248; F:pre-mRNA splicing factor activity; TAS.
 DR GO; GO:0006376; P:mRNA splice site selection; TAS.
 DR InterPro; IPR000061; Surp.
 DR Pfam; PF01805; Surp; 2.
 DR SMART; SM00648; SWAP; 2.
 DR PROSITE; PS50128; SURP; 2.
 KW Transcription regulation; RNA-binding; mRNA splicing; Repeat;

KW Nuclear protein.
 FT DOMAIN 165 168 POLY-GLU
 FT REPEAT 211 253 SURP MOTIF 1.
 FT REPEAT 459 499 SURP MOTIF 2.
 FT DOMAIN 287 290 POLY-ASP.
 FT DOMAIN 382 385 POLY-ASP.
 FT DOMAIN 413 416 POLY-PRO.
 FT DOMAIN 420 424 POLY-PRO.
 FT DOMAIN 434 440 POLY-THR.
 FT DOMAIN 451 454 POLY-PRO.
 FT DOMAIN 616 619 POLY-SER.
 FT DOMAIN 660 663 POLY-ALA.
 FT DOMAIN 754 763 POLY-LYS.
 FT DOMAIN 850 853 POLY-LYS.
 FT DOMAIN 789 951 ARG/SER-RICH (RS DOMAIN).
 SQ SEQUENCE 951 AA; 104821 MW; 02AA6C3C3CAAF24C8 CRC64;

Query Match 49.4%; Score 39; DB 1; Length 951;
 Best Local Similarity 50.0%; Pred. No. 57;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PTCQANSGRIST 14
 Db 423 PPTAFETSGATST 436

RESULT 5
 PER2_MOUSE STANDARD; PRT; 1257 AA.
 ID_PER2_MOUSE 054954;
 AC 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Period circadian protein 2 (mPER2).
 GN PER2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98088803; PubMed=9428527;
 RA Albrecht U., Sun Z.S., Bichele G., Lee C.C.;
 RT "A differential response of two putative mammalian circadian
 RT regulators, mper1 and mper2, to light";
 RL Cell 91:1055-1064(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98087121; PubMed=9427249;
 RA Shearman L.P., Zylka M.J., Weaver D.R., Kolakowski L.F. Jr.,
 RA Reppert S.M.;
 RT "Two period homologs: circadian expression and photic regulation in
 RT the suprachiasmatic nuclei";
 RL Neuron 19:1261-1269(1997).
 RN [3]
 RP REVISIONS TO 172 AND 501.
 RA Zylka M.J., Reppert S.M.;
 RA Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: CIRCADIAN REGULATOR THAT MAY ACT AS A TRANSCRIPTION
 CC FACTOR. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN TRANSCRIPTIONAL
 CC LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING INDIRECT
 CC TRANSCRIPTIONAL INHIBITION. EXPRESSION OSCILLATES WITH A 24 HR
 CC RHYTHM IN THE SUPRACHIASMATIC NUCLEUS (SCN) AND THE WHOLE EYES.
 CC OSCILLATIONS ARE MAINTAINED UNDER CONSTANT DARKNESS AND ARE
 CC RESPONSIVE TO CHANGES OF THE LIGHT/DARK CYCLES. THERE IS A 4 HOUR
 CC TIME DELAY BETWEEN PER1 AND PER2 OSCILLATIONS. THE EXPRESSION
 CC RHYTHMS APPEAR TO ORIGINATE FROM RETINA.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- TISSUE SPECIFICITY: IN THE BRAIN, HIGH EXPRESSION IN SCN DURING
 CC THE SUBJECTIVE DAY. CONSTITUTIVE EXPRESSION IN THE CORNU AMMONIS
 CC AND IN THE DENTATE GYRUS OF THE HIPPOCAMPUS. ALSO EXPRESSED IN THE
 CC PIRIFORM CORTEX AND THE GLOMERULI OF THE OLFACTORY BULB, AND AT A

CC LOWER EXTENT IN THE CEREBRAL CORTEX. NOT EXPRESSED IN THE PARS
 CC TUBERALS AND THE PURKINJE NEURONS. ALSO FOUND IN HEART, BRAIN,
 CC LUNG, LIVER, KIDNEY, AT LOW LEVEL IN SPLEEN, BUT NOT IN SKELETAL
 CC MUSCLE OR TESTIS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE SCN DURING LATE FETAL AND
 CC EARLY NEONATAL LIFE.
 CC -1- INDUCTION: BY LIGHT EXPOSURE DURING SUBJECTIVE NIGHT, BUT NOT
 CC DURING SUBJECTIVE DAY.
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS.
 CC -1- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: AF038893; AAC39942.1; -.
 CC EMBL: AF035830; AAC35592.1; -.
 CC PIR: T09493; T09493.
 CC MGD: MGI:1195265; Per2.
 CC InterPro: IPR001610; PAC.
 CC InterPro: IPR000014; PAS_domain.
 CC SMART; SM00086; PAC; 1.
 CC SMART; SM00091; PAS; 1.
 CC PROSITE; PS50112; PAS; 1.
 CC Transcription regulation; Nuclear protein; Repeat; Biological rhythms.
 CC DOMAIN 107 144 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 CC FT DOMAIN 180 246 PAS 1.
 CC FT DOMAIN 317 387 PAS 2.
 CC FT DOMAIN 336 436 PAC.
 CC FT DOMAIN 778 795 NUCLEAR LOCALIZATION SIGNAL.
 CC FT CONFLICT 445 445 P -> S (IN REF. 2).
 CC FT CONFLICT 728 728 K -> R (IN REF. 2).
 CC FT CONFLICT 1257 1257 AA; 135881 MW; 5548BAF036CF7B CRC64;
 CC SEQUENCE

Query Match 49.4%; Score 39; DB 1; Length 1257;
 Best Local Similarity 50.0%; Pred. No. 76;
 Matches 9; Conservative 2; Mismatches 3; Indels 4; Gaps 1;

QY 1 PPTCSQ----ANSGRST 14
 DB 1029 PPTCNPSPDTONSDAIST 1046

RESULT 6
 NIH2_CLOPA
 ID NIH2_CLOPA STANDARD; PRT; 272 AA.
 AC P09552;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Nitrogenase iron protein 2 (EC 1.18.6.1) (Nitrogenase component II)
 DE (Nitrogenase Fe protein 2) (Nitrogenase reductase).
 GN NIFH2.
 OS Clostridium pasteurianum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1501;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88124247; PubMed=2829127;
 RT "Wang S.-Z., Chen J.-S., Johnson J.L.;
 RT "The presence of five nifH-like sequences in Clostridium
 RT pasteurianum: sequence divergence and transcription properties.";
 RT Nucleic Acids Res. 16:439-454 (1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86168010; PubMed=3457003;

RA Chen K.C.K., Chen J.-S., Johnson J.L.;
 RT "Structural features of multiple nifH-like sequences and very biased
 RT codon usage in nitrogenase genes of Clostridium pasteurianum.";
 RL J. Bacteriol. 166:162-172 (1986).
 CC -1- FUNCTION: THE KEY ENZYMAIC REACTIONS IN NITROGEN FIXATION ARE
 CC CATALYZED BY THE NITROGENASE COMPLEX, WHICH HAS 2 COMPONENTS: THE
 CC IRON PROTEIN AND THE MOLYBDENUM-IRON PROTEIN.
 CC -1- CATALYTIC ACTIVITY: 8 reduced ferredoxin + 8 H(+) + N(2) + 16 ATP
 CC = 8 oxidized ferredoxin + 2 NH(3) + 16 ADP + 16 phosphate.
 CC -1- CORCTOR: BINDS ONE 4FE-4S CLUSTER PER DIMER.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SIMILARITY: BELONGS TO THE NIFH / BCHL / CHLL FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: X07473; CA330360.1; -.
 CC EMBL: M21537; AAA83529.1; -.
 CC PIR: S01724; S01724.
 CC HSSP: P00456; 1CP2.
 CC HAMAP; MF_00533; -; 1.
 CC InterPro; IPR005977; NifH.
 CC InterPro; IPR000392; Nitrogenaseell.
 CC InterPro; IPR000912; Nitrogenaseell.
 CC Pfam; PF00142; Fer4_NifH; 1.
 CC PRINTS; PR00091; NITROGNASIII.
 CC TIGRFAMs; TIGR01287; nifH; 1.
 CC PROSITE; PS00692; NIFH_FRXC_2; 1.
 CC PROSITE; PS00746; NIFH_FRXC_1; 1.
 CC Oxidoreductase; Nitrogen fixation; Iron-sulfur; 4Fe-4S; ATP-binding;
 CC Multigene family.
 CC NP_BIND 8 15 ATP (POTENTIAL).
 CC FT METAL 94 94 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 CC FT METAL 129 129 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 CC SEQUENCE 272 AA; 29576 MW; FB30D051EBB38FEC CRC64;

Query Match 48.1%; Score 38; DB 1; Length 272;
 Best Local Similarity 70.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PTCQANSGR 11
 DB 229 PTCQANEYR 238

RESULT 7
 NIH6_CLOPA
 ID NIH6_CLOPA STANDARD; PRT; 272 AA.
 AC P09555;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Nitrogenase iron protein 6 (EC 1.18.6.1) (Nitrogenase component II)
 DE (Nitrogenase Fe protein 6) (Nitrogenase reductase).
 GN NIFH6.
 OS Clostridium pasteurianum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1501;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88124247; PubMed=2829127;
 RT "Wang S.-Z., Chen J.-S., Johnson J.L.;
 RT "The presence of five nifH-like sequences in Clostridium
 RT pasteurianum: sequence divergence and transcription properties.";
 RT Nucleic Acids Res. 16:439-454 (1988).
 CC -1- FUNCTION: THE KEY ENZYMAIC REACTIONS IN NITROGEN FIXATION ARE
 CC CATALYZED BY THE NITROGENASE COMPLEX, WHICH HAS 2 COMPONENTS: THE
 CC IRON PROTEIN AND THE MOLYBDENUM-IRON PROTEIN.

CC -1- CATALYTIC ACTIVITY: 8 reduced ferredoxin + 8 H(+) + N(2) + 16 ATP
 CC = 8 oxidized ferredoxin + 2 NH(3) + 16 ADP + 16 phosphate.
 CC -1- COFACTOR: BINS ONE 4FE-4S CLUSTER PER DIMER.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SIMILARITY: BELONGS TO THE NIFH / BCHL / CHLL FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to 1license@1sb-sib.ch).
 CC -----
 CC EMBL: X07477; CAA30364.1; -.
 CC PIR: S02219; S02219.
 CC DR HSSP: P00456; 1CP2.
 CC DR HAMAP: MF_00533; -; 1.
 CC DR InterPro: IPR005977; NifH.
 CC DR InterPro: IPR000392; NitrogenaseEII.
 CC DR Pfam: PF00142; Fer4_NifH; 1.
 CC DR PRINTS: PR00091; NITROGENASEEII.
 CC DR TIGRFAm6: TIGR01287; nifH; 1.
 CC DR PROSITE: PS00692; NIFH_FRXC_2; 1.
 CC DR PROSITE: PS00746; NIFH_FRXC_1; 1.
 CC KW Oxidoreductase; Nitrogen fixation; Iron-sulfur; 4Fe-4S; ATP-binding;
 CC MultiGene family.
 CC FT NP_BIND 8 15 ATP (POTENTIAL).
 CC FT METAL 94 94 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 CC FT METAL 129 129 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 CC SQ SEQUENCE 272 AA; 29562 MW; DE10D05093938FEE CRC64;
 QY Query Match 48.1%; Score 38; DB 1; Length 272;
 Db Best Local Similarity 70.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 2 PTCSQANSGR 11
 229 PTCSQANSEYR 238
 RESULT 8
 ID UPAR_HUMAN STANDARD; PRT; 335 AA.
 AC Q03405; Q18876; Q15845; Q16887; Q9NYC8; Q9UDE9; Q9UEA6; Q9UM92;
 AC Q9UWV0;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Urokinase plasminogen activator surface receptor precursor (uPAR) (U-
 DE PAR) (Monocyte activation antigen Mo3) (CD87 antigen).
 GN PLAIN OR UPAR OR MO3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND SEQUENCE OF 23-33.
 RX MEDLINE=90151622; PubMed=1689240;
 RA Roldan A.L., Cobelli M.V., Masucci M.T., Behrendt N., Lund L.R.,
 RA Danos K., Appella E., Blasi F.,
 RT "Cloning and expression of the receptor for human urokinase
 RT plasminogen activator: a central molecule in cell surface, plasmin
 RT dependent proteolysis.";
 RL EMBO J. 9:467-474(1990).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=92268511; PubMed=1316922;
 RA Min H.Y., Seman R., Mizukami I.F., Watt K., Todd R.F. III, Liu D.Y.,
 RT "cDNA for Mo3, a monocyte activation antigen, encodes the human
 RT receptor for urokinase plasminogen activator.";
 RL J. Immunol. 148:3636-3642(1992).
 RN [3]

RP SEQUENCE FROM N.A. (ISOFORM 3).
 RX MEDLINE=94178507; PubMed=8131971;
 RA Bayraktutan U., Jones P.;
 RT "A novel urokinase receptor on monocyte-like macrophage cell line.";
 RL Biochem. Soc. Trans. 21:395-395(1993).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=93314820; PubMed=8392005;
 RA Pyke C., Eriksen U., Solberg H., Schnack Nielsen B., Kristensen P.,
 RA Lund L.R., Danos K.;
 RT "An alternatively spliced variant of mRNA for the human receptor for
 RT urokinase plasminogen activator.";
 RL FEBS Lett. 326:69-74(1993).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
 RC TISSUE=Placenta;
 RX MEDLINE=94325556; PubMed=8049431;
 RA Casey J.R., Petranks J.G., Kotra J., Pleonor D.E., Rose W.F.;
 RT "The structure of the urokinase-type plasminogen activator receptor
 RT gene.";
 RL Blood 84:1151-1156(1994).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Lamerdin J.E., McCready P.M., Skowronski E., Vismathan V.,
 RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnea J., Liu S.,
 RA Dargentan L., Ertler A., Christensen M., Georgescu A., Avila J., Liu S.,
 RA Altix C., Andreise T., Tranheim M., Amico-Keller G., Coefield B.,
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
 RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
 RA Olsen A.S., Carrano A.V.;
 RT "Sequence analysis of a 4.2 Mb region in 19q13.2 between AKT2 and
 RT D9S178.";
 RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.D., Usdin T.B., Toshiyuki S., Carninci P., Muller S.J.,
 RA Raha S.S., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boufield G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywicki M.I., Skalska U., Smalls D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Mera M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8]
 RP SEQUENCE OF 1-322 FROM N.A. (ISOFORM 1).
 RC TISSUE=Lung cancer;
 RA Zhu F., Jia S., He F.;
 RT "Molecular cloning of UPAR from a human giant cell lung cancer cell
 RT line.";
 RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
 RN [9]
 RP SEQUENCE OF 1-18 FROM N.A.
 RX MEDLINE=95329719; PubMed=7605992;
 RA Soravia E., Grebe A., De Luca P., Helin K., Suh T.T., Degen J.L.,
 RA Blasi F.;
 RT "A conserved TATA-less proximal promoter drives basal transcription
 RT from the urokinase-type plasminogen activator receptor gene.";
 RL Blood 86:624-635(1995).

RN PARTIAL SEQUENCE.
 RP MEDLINE=90202929; PubMed=2156852;
 RX Behrendt N., Roenne E., Ploug M., Petri T., Loeber D., Nielsen L.S.,
 RA Schleutning W.-D., Blasi F., Appella E., Danoe K.;
 RT "The human receptor for urokinase plasminogen activator. NH2-terminal
 RL amino acid sequence and glycosylation variants.";
 RN J. Biol. Chem. 265:6453-6460(1990).
 [11]
 RP SEQUENCE OF 106-116, AND CLEAVAGE BY U-PA.
 RX MEDLINE=97182578; PubMed=9030717;
 RA Hoeyer-Hansen G., Ploug M., Behrendt N., Roenne E., Danoe K.;
 RT "Cell-surface acceleration of urokinase-catalyzed receptor cleavage.";
 RL Eur. J. Biochem. 243:21-26(1997).
 [12]
 RP DISULFIDE BONDS, AND PARTIAL SEQUENCE.
 RX MEDLINE=93352546; PubMed=8394346;
 RA Ploug M., Kjaalke M., Roenne E., Weidle U., Hoeyer-Hansen G., Danoe K.;
 RT "Localization of the disulfide bonds in the NH2-terminal domain of
 RL the cellular receptor for human urokinase-type plasminogen activator.
 RT A domain structure belonging to a novel superfamily of glycolipid-
 anchored membrane proteins.";
 RL J. Biol. Chem. 268:17539-17546(1993).
 CC -1- FUNCTION: ACTS AS A RECEPTOR FOR UROKINASE PLASMINOGEN ACTIVATOR.
 CC MEDIATES THE PROTEOLYSIS-INDEPENDENT SIGNAL TRANSDUCTION.
 CC ACTIVATION EFFECTS OF U-PA. IT IS SUBJECT TO NEGATIVE-FEEDBACK
 CC REGULATION BY U-PA WHICH CLEAVES IT INTO AN INACTIVE FORM.
 CC -1- SUBUNIT: Monomer (Probable).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (isoform 1). Secreted (isoform 2).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Name=1; Synonyms=uPAR1, GPI-anchored;
 CC IsoId=Q03405-1; Sequence=Displayed;
 CC Name=2; Synonyms=uPAR2, Secreted;
 CC IsoId=Q03405-2; Sequence=VSP_006715;
 CC Name=3;
 CC IsoId=Q03405-3; Sequence=VSP_006714;
 CC -1- SIMILARITY: Contains 3 uPAR/Ly6 domains.
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD87 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd87.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, X51675; CAA35981.1; -;
 DR EMBL, M83246; AAA59862.1; -;
 DR EMBL, X74039; CAA52191.1; -;
 DR EMBL, U09346; AAA17979.1; -;
 DR EMBL, U09347; AAA17978.1; -;
 DR EMBL, U09937; AAB60690.1; -;
 DR EMBL, U09931; AAB60690.1; JOINED.
 DR EMBL, U09932; AAB60690.1; JOINED.
 DR EMBL, U09933; AAB60690.1; JOINED.
 DR EMBL, U09935; AAB60690.1; JOINED.
 DR EMBL, U09936; AAB60690.1; JOINED.
 DR EMBL, U08839; AAB60333.1; -;
 DR EMBL, AC005525; AAC32733.1; -;
 DR EMBL, AC006953; AAD17388.1; -;
 DR EMBL, AC006953; AAD17387.1; -;
 DR EMBL, BC002788; AAH02788.1; -;
 DR EMBL, AF257789; AAF71751.1; -;
 DR EMBL, S78532; AAD14289.1; -;
 DR EMBL, A18757; CAA01421.1; -;
 DR PIR, S39495; S39495.
 DR PIR, S39495; S39495.
 DR Genew; HGNC:9053; PLANR.

DR MIM, 173391; -;
 DR GO; GO:0005866; C:plasma membrane; TAB.
 DR GO; GO:001025; F:GPI-anchored membrane-bound receptor; TAB.
 DR GO; GO:0007596; F:blood coagulation; TAB.
 DR GO; GO:0006935; P:chemotaxis; TAB.
 DR InterPro; IPR001526; Ly6 UPAR.
 DR InterPro; IPR003631; U-PA.
 DR Pfam; PF00021; UPAR Ly6; 2.
 DR ProDom; PD007582; U-PA; 1.
 DR SMART; SM00134; Ly6 UPAR; 3.
 DR PROSITE; PS00983; Ly6 UPAR; 3.
 DR Receptor; Signal; Glycoprotein; GPI-anchor; Repeat;
 KW Alternative splicing; Polymorphism.
 FT SIGNAL 1 22
 FT CHAIN 23 7305 UROKINASE PLASMINOGEN ACTIVATOR SURFACE
 FT PROPEP 7306 335 RECEPTOR.
 FT DOMAIN 23 114 REMOVED IN MATURE FORM (POTENTIAL).
 FT DOMAIN 115 213 UPAR/Ly6 1.
 FT DOMAIN 214 305 UPAR/Ly6 2.
 FT CARBOHYD 74 74 UPAR/Ly6 3.
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 222 222 N-LINKED (GLCNAC. . .) (POTENTIAL).
 Query Match 48.1%; Score 38; DB 1; Length 335;
 Best Local Similarity 63.6%; Pred No. 29;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 4 CSQANSGRIST 14
 Db 98 CNOGNSGRAVT 108
 ID Y479 MYCTU STANDARD; PRT; 348 AA.
 AC 01145;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein Rv0479c.
 GN Rv0479C OR M70497 OR M7C2059.05C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxId=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigler K., Gae S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultson J.E., Taylor K., Whitehead S., Barrill B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RT Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 277162; CAB00940.1; -
CC EMBL; AB006951; AAK44720.1; -
CC DR PIR; B70743; B70743.
CC DR TIGR; MT0497; -
CC DR Tuberculese; Rv0479c; -
CC KM Hypothetical protein; Transmembrane; Complete proteome.
CC FT TRANSMEM 111 131 POTENTIAL.
CC FT TRANSMEM 235 255 POTENTIAL.
CC SQ SEQUENCE 348 AA; 37048 MW; 9EBB7A8CF9E176C1 CRC64;
CC -----
Query Match 48.1%; Score 38; DB 1; Length 348;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
OY 1 PPTCSQANSGR1 12
DB 26 PPTASSEASTGR1 37
-----
RESULT 10
MT8 THETH STANDARD; PRT; 428 AA.
AC P29749;
AD 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Modification methylase TthHB81 (EC 2.1.1.72) (Adenine-specific
DE methyltransferase TthHB81) (M.TthHB81).
GN TTHHB81M.
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=274;
OX [1]
RX SEQUENCE FROM N.A.
RC STRAIN=HB8 / ATCC 27634;
RA MEDLINE=92201699; PubMed=1339363;
RA Barany F., Danzitz M., Zebala J., Mayer A.;
RT "Cloning and sequencing of genes encoding the TthHB81 restriction and
RT modification enzymes: comparison with the isoschizomeric TaqI
RT enzymes.";
RL Gene 112:3-12(1992).
CC -1- FUNCTION: THIS METHYLASE RECOGNIZES THE DOUBLE-STRANDED SEQUENCE
CC TCGA, CAUSES SPECIFIC METHYLATION ON A-4 ON BOTH STRANDS AND
CC PROTECTS THE DNA FROM CLEAVAGE BY THE TTHHB81 ENDONUCLEASE.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-
CC adenosyl-L-homocysteine + DNA 6-methylaminopurine.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M74795; AAA27488.1; -
CC HSSP; P14385; 2ADM.
CC DR REBASE; 3523; M.TthHB81.
CC DR InterPro; IPR002296; N12M6_mtfase.
CC DR InterPro; IPR002052; N6_Mtase.
CC DR InterPro; IPR000051; SAM_bind.
CC DR PRINTS; PRO0507; N12M6MTFRASE.
CC DR PROSITE; PS00092; N6_MTASE; 1.

```

```

KM Transferase; Methyltransferase; Restriction system.
SQ SEQUENCE 428 AA; 48317 MW; 4CDB8B1959EFP6 CRC64;
-----
Query Match 48.1%; Score 38; DB 1; Length 428;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
OY 1 PPTCSQANSGR1ST 14
DB 8 PPTASGRSLGRVET 21
-----
RESULT 11
ANA DROME STANDARD; PRT; 474 AA.
AC Q26307; Q9V524;
AD 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Anachronism protein precursor.
GN ANA OR CG8084.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
OX [1]
RX SEQUENCE FROM N.A.
RC TISSUE=Eye; imaginal disk;
RX MEDLINE=93327423; PubMed=7916657;
RA Ebens A.J., Garren H., Cheyette B.N.R., Zipursky S.L.;
RT "The Drosophila anachronism locus: a glycoprotein secreted by glia
RT inhibits neuroblast proliferation.";
RL Cell 74:15-27(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster A., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali B., Kalush F., Karpen G.H., Ke Z., Kenton J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasio P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Martel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Werklow G., Millett N.V., Mobarry C., Morris C., Moshirizadeh A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Pacle J.M.,
RA Palazkoto M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Scapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

```

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 CC Science 287:2185-2195(2000).
 CC -1- FUNCTION: NEGATIVELY REGULATES PROLIFERATION OF NEURONAL PRECURSOR
 CC CELLS, THEREBY CONTROLLING THE TIMING OF POSTEMBRYONIC
 CC NEUROGENESIS;
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN SOME GLIAL CELLS AND SECRETED.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; S63815; AAB27582.1; -;
 CC DR EMBL; AE003834; AAF58998.1; -;
 CC DR FlyBase; FBgn0011746; ana.
 CC KM Glycoprotein; Signal; Developmental protein; Neurogenesis.
 CC FT SIGNAL 1 33 POTENTIAL.
 CC FT CHAIN 34 474 ANACHRONISM PROTEIN.
 CC FT DOMAIN 453 462 POLY-HIS.
 CC FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 342 342 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CONFLICT 388 388 A -> G (IN REF. 1).
 CC SQ SEQUENCE 474 AA; 54015 MW; FE8A43042BB595A1 CRC64;
 CC -----
 CC Query Match 48.1%; Score 38; DB 1; Length 474;
 CC Best Local Similarity 75.0%; Pred. No. 42;
 CC Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC QY 1 PPTCSOAN 8
 CC DB 302 PPTCDQAD 309
 CC -----
 CC RESULT 12
 CC COEL RAT STANDARD; PRT; 584 AA.
 CC ID COEL RAT
 CC AC 063358;
 CC DT 30-MAY-2000 (Rel. 39, Created)
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Transcription factor COEL (OE-1) (O/E-1) (Early B-cell factor)
 CC DE (Olfactory neuronal transcription factor) (Olf-1).
 CC GN EBF OR EBF1 OR COEL OR OLF1.
 CC OS Rattus norvegicus (Rat).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC OX NCBI_Taxid=10116;
 CC RN [1]
 CC RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 CC RX MEDLINE=93309542; PubMed=8321284;
 CC RT Wang M.M., Reed R.R.;
 CC RT "Molecular cloning of the olfactory neuronal transcription factor
 CC RT Olf-1 by genetic selection in yeast.";
 CC RT Nature 364:121-126(1993).
 CC CC -1- FUNCTION: Transcriptional activator which recognizes variations
 CC of the palindromic sequence 5'-ATTCCNNGGAATT-3'.
 CC CC -1- SUBUNIT: Forms either a homodimer or a heterodimer with a related
 CC family member.
 CC CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC CC -1- TISSUE SPECIFICITY: Expressed exclusively in olfactory receptor
 CC neurons and their precursors.
 CC CC -1- SIMILARITY: BELONGS TO THE COE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L24051; AAA41759.1; ALT INIT.
 CC DR InterPro; IPR001092; HLH basic.
 CC DR InterPro; IPR002909; IPT-TIG.
 CC DR InterPro; IPR003523; TF_COE.
 CC DR Pfam; PF01833; TIG; 1.
 CC DR SMART; SM00353; HLH; 1.
 CC DR SMART; SM00429; IPT; 1.
 CC DR PROSITE; PS01345; COE; 1.
 CC KM Transcription regulation; DNA-binding; Nuclear protein; Activator;
 CC FT Developmental protein; Zinc-finger.
 CC FT ZN_FING 151 170 CS-TYPE (POTENTIAL).
 CC FT DOMAIN 455 543 SSR/THR-PRO-RICH.
 CC SQ SEQUENCE 584 AA; 63650 MW; 1E903A9801BCEFA CRC64;
 CC -----
 CC Query Match 48.1%; Score 38; DB 1; Length 584;
 CC Best Local Similarity 33.3%; Pred. No. 52;
 CC Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 CC -----
 CC QY 1 PPTCSOANSGRISTL 15
 CC DB 562 PPTCTNGNSLOAI 576
 CC -----
 CC RESULT 13
 CC COEL HUMAN STANDARD; PRT; 586 AA.
 CC ID COEL HUMAN
 CC AC 09UW73;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
 CC DE Transcription factor COEL (OE-1) (O/E-1) (Early B-cell factor)
 CC DE (Fragment).
 CC GN EBF OR EBF1 OR COEL.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CC OX NCBI_Taxid=9606;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=20400007; PubMed=10942392;
 CC RA Gistler R., Jacobsen S.E., Sigvardsson M.;
 CC RT "Cloning of human early B-cell factor and identification of target
 CC RT genes suggest a conserved role in B-cell development in man and
 CC RT mouse.";
 CC RT Blood 96:1457-1464(2000).
 CC CC -1- FUNCTION: Transcriptional activator which recognizes variations of
 CC the palindromic sequence 5'-ATTCCNNGGAATT-3'. (By similarity).
 CC CC -1- SUBUNIT: Forms either a homodimer or a heterodimer with a related
 CC family member (By similarity).
 CC CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC CC -1- SIMILARITY: BELONGS TO THE COE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF208502; AAF19643.1; -;
 CC DR Genew; HGNC:3126; EBF.
 CC DR MIM; 164343;
 CC DR InterPro; IPR001092; HLH basic.
 CC DR InterPro; IPR002909; IPT-TIG.
 CC DR InterPro; IPR003523; TF_COE.

```

DR Pfam: PF01833; TIG: 1.
DR SMART, SM00353; HLH; 1.
DR PROSITE, PS00429; IPT; 1.
DR PROSITE, PS01345; COE; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; Activator;
KW Developmental protein; Zinc-finger.
FT NON TER 1 1
FT ZN FING 146 165 CS-TYPE (POTENTIAL).
FT DOMAIN 457 545 SER/THR/PRO-RICH.
SQ SEQUENCE 586 AA; 63845 MW; D3596C26FEC72CF6 CRC64;

Query Match 48.1%; Score 38; DB 1; Length 586;
Best Local Similarity 33.3%; Pred. No. 52;
Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 PPTCSQANSGRISTL 15
Db 564 PPTCTSTNGNSLOAI 578

RESULT 14
ID COE1_MOUSE STANDARD; PRT; 591 AA.
AC 007802;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transcription factor COE1 (OE-1) (Early B-cell factor).
GN EBF OR EBF1 OR COE1.
OS Mus musculus (Mouse).
OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
(1)
RP SEQUENCE FROM N.A., AND SEQUENCE OF 87-99; 122-129; 342-358 AND
RP 382-390.
RC TISSUE=Lymphoid;
RX MEDLINE=93359462; PubMed=8491377;
RA Hagman J., Belanger C., Travis A., Turck W., Groschedl R.;
RT "Cloning and functional characterization of early B-cell factor, a
RT regulator of lymphocyte-specific gene expression.";
RT Gene Dev. 7:760-773(1993).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93268288; PubMed=8497258;
RA Travis A., Hagman J., Hwang L., Groschedl R.;
RT "Purification of early-B-cell factor and characterization of its
RT DNA-binding specificity.";
RT Mol. Cell. Biol. 13:3392-3400(1993).
RN [3]
RP SUBUNIT, AND ALTERNATIVE SPLICING.
RC STRAIN=CD-1; TISSUE=Embryo;
RX MEDLINE=97296547; PubMed=9151732;
RA Wang S.S., Tsai R.Y.L., Reed R.R.;
RT "The characterization of the Olf-1/EBF-like HLH transcription factor
RT family: implications in olfactory gene regulation and neuronal
RT development.";
RT J. Neurosci. 17:4149-4158(1997).
RL
CC -1- FUNCTION: Transcriptional activator which recognizes variations of
CC the palindromic sequence 5'-ATTCGNGGGAATT-3'.
CC -1- SUBUNIT: Forms either a homodimer or a heterodimer with a related
CC family member.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long; Synonyms=8;
CC Name=Short; Synonyms=0;
CC IsoId=Q07802-1; Sequence=displayed;
CC IsoId=Q07802-2; Sequence=VSP_001112.
CC -1- TISSUE SPECIFICITY: Expressed at high levels in early B-cells,
CC spleen, lymph node and adipose tissue, and at low levels in heart,
CC brain, skeletal muscle and kidney. In adult expressed in olfactory
CC epithelium, in spleen, and at a lesser extent in Purkinje cells of

```

```

CC the cerebellum, heart, kidney, lung, thymus and testis. In embryo
CC expressed in dorsal thalamus and epithalamus, at a lower level in
CC mesencephalon and in the caudal rhombencephalon. In the
CC postmitotic cells of developing retina, highly in developing
CC spinal cord, dorsal root ganglia, trigeminal ganglia and in
CC gliospharyngeal nerve ganglia, in developing inner ear.
CC -1- DEVELOPMENTAL STAGE: First detected at 9.5 dpc.
CC -1- SIMILARITY: BELONGS TO THE COE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
CC EMBL, L12147; AAA37533.1; -.
CC PIR, A40684; A40684.
CC TRANSFAC, T01112; -.
CC MGD, MGI:95275; EBF1.
CC DR GO; GO:0005515; F:protein binding activity; IPT.
CC DR GO; GO:0045941; P:positive regulation of transcription; IDA.
CC DR InterPro; IPR001092; HLH basic.
CC DR InterPro; IPR002909; IPT_TIG.
CC DR InterPro; IPR003523; TF_COE.
CC DR Pfam; PF01833; TIG; 1.
CC DR SMART; SM00353; HLH; 1.
CC DR SMART; SM00429; IPT; 1.
CC DR PROSITE; PS01345; COE; 1.
CC DR Transcription regulation; DNA-binding; Nuclear protein; Activator;
CC Developmental protein; Alternative splicing; Zinc-finger.
CC ZN FING 151 170 CS-TYPE (POTENTIAL).
CC FT DOMAIN 462 550 SER/THR/PRO-RICH.
CC FT VARSPLIC 252 259 Missing (in isoform Short).
CC FT FTID=VSP_001112.
CC FT CONFLICT 87 87 T -> F (IN REF. 1; AA SEQUENCE).
CC FT CONFLICT 89 89 F -> L (IN REF. 1; AA SEQUENCE).
CC SQ SEQUENCE 591 AA; 64464 MW; E47797B6FCE5071 CRC64;

Query Match 48.1%; Score 38; DB 1; Length 591;
Best Local Similarity 33.3%; Pred. No. 52;
Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 PPTCSQANSGRISTL 15
Db 569 PPTCTSTNGNSLOAI 583

RESULT 15
ID CCAF_HUMAN STANDARD; PRT; 1966 AA.
AC 060840; O43901;
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Voltage-dependent L-type calcium channel alpha-1F subunit.
GN CACNA1F OR CACNA1F.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
(1)
RP SEQUENCE FROM N.A., AND VARIANTS CSNB2 ASP-369; GLN-508; TRP-1049 AND
RP HIS-1364.
RC TISSUE=Retina;
RX MEDLINE=98324776; PubMed=9662399;
RX Strom T.M., Nakamura G., Apfelstedt-Sylla E., Hellebrand H.,
RX Lorenz B., Weber B.H.F., Wulz K., Gutwillinger N., Rueher K.,
RX Drescher B., Sauer C., Zrenner E., Meltlinger T., Rosenthal A.,
RX Weidl A.;
RT "An L-type calcium-channel gene mutated in incomplete X-linked
RT congenital stationary night blindness.";

```

RL Nat. Genet. 19:260-263 (1998).
 RN (2)
 RP SEQUENCE FROM N.A.
 RA Blechschmidt K., Nyakatura G., Strom T.M., Drescher B., Menzel U.,
 RA Meindl A., Rosenblatt A.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBD databases.
 RN [3]
 RP SEQUENCE OF 1200-1966 FROM N.A.
 RX MEDLINE=98008923; PubMed=9344658;
 RA Fisher S.E., Ciccodicola A., Tanaka K., Curci A., Desicic S.,
 RA D'Urso M., Craig I.W.;
 RT "Sequence-based exon prediction around the synaptophysin locus reveals
 RT a gene-rich area containing novel genes in human proximal Xp.";
 RL Genomics 45:340-347 (1997).
 CC -1- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE
 CC ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
 CC IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
 CC CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,
 CC CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOTFORM ALPHA-1F
 CC GIVES RISE TO L-TYPE CALCIUM CURRENTS. LONG-LASTING (L-TYPE)
 CC CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA)
 CC GROUP. THEY ARE BLOCKED BY DIHYDROPYRIDINES (DHP),
 CC PHENYLTALKALAMINES, BENZOTHAZEPINES, AND BY OMEGA-AGATOXIN-111A
 CC (OMEGA-AGA-111A). THEY ARE HOWEVER INSENSITIVE TO OMEGA-CONOTOXIN-
 CC GYIA (OMEGA-CTX-GYIA) AND OMEGA-AGATOXIN-1YA (OMEGA-AGA-1YA).
 CC -1- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
 CC COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
 CC IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-
 CC FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
 CC SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
 CC CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
 CC LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSION IN SKELETAL MUSCLE AND RETINA.
 CC -1- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
 CC HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
 CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
 CC PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
 CC SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
 CC -1- DISEASE: Defects in CACNA1F are the cause of incomplete X-linked
 CC congenital stationary night blindness type 2 (CSNB2) [MIM:300071],
 CC a nonprogressive retinal disorder characterized by decreased
 CC visual acuity and loss of night vision.
 CC -1- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 CC EMBL; AJ006216; CAA06916.1; -
 DR EMBL; AJ224874; CAA12175.1; -
 DR EMBL; AF235097; -; NOT ANNOTATED_CDS.
 DR EMBL; U93305; AAB92359.1; -
 DR GeneW; HGNC:1393; CACNA1F.
 DR MIM; 300110; -
 DR MIM; 300071; -
 DR GO; GO:0005891; C:voltage-gated calcium channel complex; TAS.
 DR GO; GO:0005895; F:voltage-gated calcium channel activity; TAS.
 DR GO; GO:0006833; P:small molecule transport; TAS.
 DR GO; GO:0007268; P:synaptic transmission; TAS.
 DR GO; GO:0007601; P:vision; TAS.
 DR InterPro; IPR001682; Ca/Na_pore.
 DR InterPro; IPR002077; Ca_channel.
 DR InterPro; IPR002121; Cat_channel_Tryp.
 DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR005446; LVDCa1phal.
 DR InterPro; IPR005820; M-channel_nlg.
 DR Pfam; PF00520; Ion_trans; 4.

DR	PRINTS: PRO0167; CACHANNEL.
KM	PRINTS: PRO1630; LVDCCALPHAI.
KW	Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
KV	Calcium channel; Glycoprotein; Repeat; Multisene family;
KX	Calcium-binding; Phosphorylation; Disease mutation; Vision.
FT	REPEAT 79 375 I.
FT	REPEAT 504 750 II.
FT	REPEAT 847 1129 III.
FT	REPEAT 1166 1433 IV.
FT	DOMAIN 1 92 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 93 111 S1 OF REPEAT I (POTENTIAL).
FT	DOMAIN 112 129 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 130 149 S2 OF REPEAT I (POTENTIAL).
FT	DOMAIN 150 161 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 162 180 S3 OF REPEAT I (POTENTIAL).
FT	DOMAIN 181 201 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 202 220 S4 OF REPEAT I (POTENTIAL).
FT	DOMAIN 221 239 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 240 259 S5 OF REPEAT I (POTENTIAL).
FT	DOMAIN 260 347 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 348 372 S6 OF REPEAT I (POTENTIAL).
FT	DOMAIN 373 518 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 519 553 S1 OF REPEAT II (POTENTIAL).
FT	DOMAIN 539 553 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 554 572 S2 OF REPEAT II (POTENTIAL).
FT	DOMAIN 573 580 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 581 599 S3 OF REPEAT II (POTENTIAL).
FT	DOMAIN 600 609 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 610 628 S4 OF REPEAT II (POTENTIAL).
FT	DOMAIN 629 647 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 648 722 S5 OF REPEAT II (POTENTIAL).
FT	DOMAIN 669 768 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 723 747 S6 OF REPEAT II (POTENTIAL).
FT	DOMAIN 748 860 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 861 879 S1 OF REPEAT III (POTENTIAL).
FT	DOMAIN 880 895 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 896 915 S2 OF REPEAT III (POTENTIAL).
FT	DOMAIN 916 927 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 928 946 S3 OF REPEAT III (POTENTIAL).
FT	DOMAIN 947 952 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 953 972 S4 OF REPEAT III (POTENTIAL).
FT	DOMAIN 973 991 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 992 1011 S5 OF REPEAT III (POTENTIAL).
FT	DOMAIN 1012 1101 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 1102 1122 S6 OF REPEAT III (POTENTIAL).
FT	DOMAIN 1123 1179 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 1180 1198 S1 OF REPEAT IV (POTENTIAL).
FT	DOMAIN 1199 1213 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 1214 1233 S2 OF REPEAT IV (POTENTIAL).
FT	DOMAIN 1234 1240 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 1241 1262 S3 OF REPEAT IV (POTENTIAL).
FT	DOMAIN 1263 1279 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 1280 1299 S4 OF REPEAT IV (POTENTIAL).
FT	DOMAIN 1300 1318 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 1319 1338 S5 OF REPEAT IV (POTENTIAL).
FT	DOMAIN 1339 1405 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 1406 1430 S6 OF REPEAT IV (POTENTIAL).
FT	DOMAIN 1431 1966 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 648 654 POLY-LEU.
FT	DOMAIN 783 788 POLY-GLU.
FT	DOMAIN 798 814 POLY-GLU.
FT	DOMAIN 1110 1113 POLY-TLE.
FT	DOMAIN 1629 1634 POLY-GLU.
FT	DOMAIN 395 412 BINDING TO THE BETA SUBUNIT (BY SIMILARITY).
FT	SITE 330 330 CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).
FT	SITE 700 700 CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).
FT	SITE 1075 1075 CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).
FT	SITE 1372 1372 CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).


```

FT BINDING 1049 1139 DIHYDROPYRIDINES (BY SIMILARITY).
FT BINDING 1386 1452 DIHYDROPYRIDINES (BY SIMILARITY).
FT BINDING 1398 1441 PHENYLALKYLAMINES (BY SIMILARITY).
FT MOD_RES 1441 1441 PHOSPHORYLATION (BY PKA) (POTENTIAL).
FT CA_BIND 1459 1470 BY SIMILARITY.
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 369 369 G -> D (in CSNB2).
FT VARIANT 369 369 /FTid=VAR_001504.
FT VARIANT 508 508 R -> Q (in CSNB2).
FT VARIANT 1049 1049 /FTid=VAR_001505.
FT VARIANT 1049 1049 R -> W (in CSNB2).
FT VARIANT 1364 1364 /FTid=VAR_001506.
FT VARIANT 1364 1364 L -> H (in CSNB2).
FT CONFLICT 1225 1225 /FTid=VAR_001507.
FT CONFLICT 1265 1271 E -> V (in REF. 3).
FT CONFLICT 1652 1767 MISSING (in REF. 3).
FT CONFLICT 1849 1849 MISSING (in REF. 3).
FT CONFLICT 1849 1849 A -> G (in REF. 3).
SQ SEQUENCE 1966 AA, 219494 MW, PEB47E19FA57E31D CRC64,

```

Query Match 48.1%; Score 38; DB 1; Length 1966;

Best Local Similarity 46.7%; Pred. No. 1.8e+02; Mismatches 6; Indels 0; Gaps 0;

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 PPTCSQANSGRISTL 15

DB 280 PSPCASSGSGRACTL 294

Search completed: November 26, 2003, 12:31:19
Job time : 8.22892 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:25:40 ; Search time 13.9157 Seconds
(without alignments)
103.662 Million cell updates/sec

Title: US-09-230-111C-14

Perfect score: 79

Sequence: 1 PPTCSQANSGRSTL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	55.7	409	2	T03788
2	43	54.4	699	2	D70667
3	42	53.2	1257	2	T13957
4	42	53.2	1777	2	T13469
5	41	51.9	133	2	T08792
6	41	51.9	277	2	T12853
7	40	50.6	304	2	T29421
8	40	50.6	419	2	D85334
9	40	50.6	419	2	T04530
10	40	50.6	474	2	S66480
11	40	50.6	557	2	JCS229
12	39	49.4	203	2	S32799
13	39	49.4	508	1	KRSHL2
14	39	49.4	749	2	E95889
15	39	49.4	942	2	JC2129
16	39	49.4	1257	2	T09493
17	39	49.4	1559	2	T10535
18	38	48.1	84	2	P90329
19	38	48.1	110	2	T26729
20	38	48.1	134	2	D97582
21	38	48.1	134	2	AD2803
22	38	48.1	244	1	A64843
23	38	48.1	272	2	S01724
24	38	48.1	272	2	S02219
25	38	48.1	281	2	S39495
26	38	48.1	303	2	P90301
27	38	48.1	321	2	C90402
28	38	48.1	321	2	D90326
29	38	48.1	335	2	A39743

30	38	48.1	348	2	B70743	hypothetical prote
31	38	48.1	360	2	T18529	proctein phosphatase
32	38	48.1	366	2	P90474	hypothetical prote
33	38	48.1	428	2	JH0634	site-specific DNA-
34	38	48.1	439	2	P96592	probable zinc fing
35	38	48.1	474	2	A40721	neuroblast prolif
36	38	48.1	570	2	S35069	transcription fact
37	38	48.1	591	2	A40684	early B-cell facto
38	38	48.1	779	2	S64993	hypothetical prote
39	38	48.1	1160	2	P88369	protein unc-52 (1m
40	38	48.1	1365	2	T45031	hypothetical prote
41	38	48.1	2295	2	C88369	protein unc-52 (1m
42	38	48.1	3375	2	T19821	hypothetical prote
43	37.5	47.5	115	2	T20559	hypothetical prote
44	37	46.8	79	2	AG2767	hypothetical prote
45	37	46.8	162	2	D96581	hypothetical prote

ALIGNMENTS

RESULT 1
T03788
laccase (EC 1.10.3.2) - common tobacco (fragment)
N/Alternate names: diphenol oxidase; urushiol oxidase
C/Species: Nicotiana tabacum (common tobacco)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C/Accession: T03788
R/Kiefer-Meyer, M.; Gomord, V.; O'Connell, A.; Halpin, C.; Faye, L.
Gene 178, 205-207, 1996
A/Title: Cloning and sequence analysis of laccase-encoding cDNA clones from tobacco.
A/Reference number: JCS229; MUID:97080572; PMID:8921917
A/Accession: T03788
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-409 <KIR>
A/Cross-references: EMBL:U45243; NID:G1685090; PIDN:AAC49538.1; PID:G1685091
C/Superfamily: laccase
C/Keywords: copper; oxidoreductase

Query Match 55.7% Score 44; DB 2; Length 409;
Best Local Similarity 63.6% Pred. No. 13;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PPTCSQANSGR 12
DB 218 PPTCSQANSRV 228

RESULT 2
D70667
hypothetical protein Rv1668 - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C/Accession: D70667

R/Coile, S.T.; Broech, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: AT0500; MUID:98295887; PMID:9634230

A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Accession: D70667

A/Molecule type: DNA
A/Residues: 1-699 <COL>
A/Cross-references: GB:Z83859; GB:AL123456; NID:G3261678; PIDN:CAB06119.1; PID:e290717;

A/Experimental source: strain H37RV
C/Genetics:
A/Genes: Rv1668

Query Match 54.4% Score 43; DB 2; Length 699;
Best Local Similarity 40.0% Pred. No. 31;

Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 1 PPTCSQANSGRISTL 15
|||:|:|:|:
DB 530 PPLCALANDGNLAI 544

RESULT 3

T13957
period protein PER2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T13957
R.Sakamoto, K.; Nagase, T.; Fukui, H.; Horikawa, K.; Okada, T.; Tanaka, H.; Sato, K.; Mi-
J. Biol. Chem. 273: 27039-27042, 1998
A:Title: Multitissue circadian expression of rat period homologue (rper2) mRNA is govern
A:Reference number: Z17833; MIMD:98438460; PMID:9765215
A:Accession: T13957
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1257 <SAK>
A:Cross-references: EMBL:AB016532; NID:g3845582; PIDN:BAA34187.1; PID:g3845583
A:Experimental source: strain Sprague-Dawley; brain; clone RG1528
C:Genetics:
A:Gene: rper2
C:Keywords: circadian rhythm

Query Match 53.2%; Score 42; DB 2; Length 1257;
Best Local Similarity 55.6%; Pred. No. 79;
Matches 10; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 1 PPTCSO---ANSGRIST 14
|||:|:|:|:
DB 1029 PPTCSFSPSTQNSDAIST 1046

RESULT 4

T34369
hypothetical protein T19D12.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34369
R.Favell, A.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid T19D12.
A:Reference number: Z21513
A:Accession: T34369
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1777 <FAV>
A:Cross-references: EMBL:U41263; PIDN:AAC24428.1; GSPDB:GN00020; CESP:T19D12.1
A:Experimental source: strain Bristol N2; clone T19D12
C:Genetics:
A:Gene: CESP:T19D12.1
A:Map position: 2

A:Introns: 36/1; 134/2; 180/1; 622/3; 691/2; 754/1; 1111/2; 1174/1; 1271/3; 1322/2; 1681
Query Match 53.2%; Score 42; DB 2; Length 1777;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 PTCQANSGRISTL 15
|||:|:|:|:
DB 1659 PTCASPYSGKIATV 1672

RESULT 5

T08792
hypothetical protein DKFZ586E1422.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999
C:Accession: T08792
R.Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, March 1999
A:Reference number: Z16473
A:Accession: T08792
A:Molecule type: mRNA
A:Residues: 1-133 <KOE>
A:Cross-references: EMBL:AL050170
A:Experimental source: adult uterus; clone DKFZ586E1422
C:Genetics:
A:Note: DKFZ586E1422.1

Query Match 51.9%; Score 41; DB 2; Length 133;
Best Local Similarity 63.6%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PTCQANSGR 12
|||:|:|:|:
DB 115 PACSQAGSAGV 125

RESULT 6

T32853
hypothetical protein K05F6.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32853
R.Du, Z.; Goela, D.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid K05F6.
A:Reference number: Z21233
A:Accession: T32853
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-277 <DU>
A:Cross-references: EMBL:AF040653; PIDN:AAB95023.1; GSPDB:GN00020; CESP:K05F6.10
A:Experimental source: strain Bristol N2; clone K05F6
C:Genetics:
A:Gene: CESP:K05F6.10
A:Map position: 2
A:Introns: 139/1; 201/2; 232/1

Query Match 51.9%; Score 41; DB 2; Length 277;
Best Local Similarity 57.1%; Pred. No. 28;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 PPTCSQANSGRIST 14
|||:|:|:|:
DB 51 PPTAKANFGRBGT 64

RESULT 7

T29421
hypothetical protein SC3A7.03c - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T29421
R.Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1998
A:Reference number: Z20619
A:Accession: T29421
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-304 <PAR>
A:Cross-references: EMBL:AL031155; NID:e1313489; PID:e1313492; PIDN:CAA20066.1
C:Genetics:
A:Note: SC3A7.03c

Query Match 50.6%; Score 40; DB 2; Length 304;
Best Local Similarity 54.5%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PPTCSQANSGR 11
|||:|:|:|:
DB 115 PPAQNSRRR 125

RESULT 8

D85334

Hypothetical protein AT4g28690 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001

C:Accession: D85334

R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring

Nucleus 402, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; MUID:20083488; PMID:10617198

A:Accession: D85334

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-419 <STO>

A:Cross-references: GB:NC_001268; NID:97269724; PIDN:CAB81457.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4g28690

A:Map position: 4

Query Match

Best Local Similarity 50.6%; Score 40; DB 2; Length 419;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 PPTCSQANSGRIST 14

Db 315 PPVASQGNARLVT 328

RESULT 9

T04530

Hypothetical protein F16A16.200 - Arabidopsis thaliana (fragment)

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999

C:Accession: T04530

R:Bayan, M.; Brandt, P.; Dose, S.; Jarke, D.; Scharfe, M.; Schon, O.; Hohelsel, J.; Mewe

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15376

A:Accession: T04530

A:Molecule type: DNA

A:Residues: 1-419 <BEV>

A:Cross-references: EMBL:AL035353

A:Experimental source: cultivar Columbia; BAC clone F16A16

C:Genetics:

A:Map position: 4

A:Note: Intron positions not resolved

A:Note: F16A16.200

Query Match

Best Local Similarity 50.6%; Score 40; DB 2; Length 419;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 PPTCSQANSGRIST 14

Db 315 PPVASQGNARLVT 328

RESULT 10

S66480

Carbon catabolite repressor MIG1 - yeast (Kluyveromyces marxianus var. lactis)

N:Alternate names: finger protein MIG1

C:Species: Kluyveromyces marxianus var. lactis, Candida sphaerica

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 29-Oct-1999

C:Accession: S66480

R:Caesari, J.P.; Georls, I.; Oeseling, J.; Rone, H.; Vandenhaute, J.

PDBS Lett. 371, 191-194, 1995

A:Title: The MIG1 repressor from Kluyveromyces lactis: cloning, sequencing and function

A:Reference number: S66480; MUID:95402208; PMID:7672126

A:Accession: S66480

A:Molecule type: DNA

A:Residues: 1-474 <CAS>

A:Cross-references: EMBL:Z50017; NID:g1109764; PIDN:CAA90320.1; PID:el87972; PID:g110976

A:Experimental source: strain CBS 2359

A:Note: the source is designated as Kluyveromyces lactis

C:Genetics:

A:Gene: MIG1

A:Keywords: DNA binding; transcription regulation; zinc finger

F:28-78/Region: zinc finger CCH motif

Query Match

Best Local Similarity 50.6%; Score 40; DB 2; Length 474;

Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PPTCSQANSGRIST 15

Db 178 PPIRSQNSGNIDDL 192

RESULT 11

JC5229

laccase (EC 1.10.3.2) precursor - common tobacco

N:Alternate names: p-diphenol:O2 oxidoreductase; urushiol oxidase

C:Species: Nicotiana tabacum (common tobacco)

C:Date: 13-Mar-1997 #sequence_revision 13-Mar-1997 #text_change 11-Jun-1999

C:Accession: JC5229

R:Kiefer-Meyer, M.; Gomord, V.; O'Connell, A.; Halpin, C.; Faye, L.

Gene 178, 205-207, 1996

A:Title: Cloning and sequence analysis of laccase-encoding cDNA clones from tobacco.

A:Reference number: JC5229; MUID:97080572; PMID:8921917

A:Accession: JC5229

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-557 <KIE>

A:Cross-references: GB:U43542; NID:g1685086; PIDN:AAC49536.1; PID:g1685087

C:Comment: This enzyme oxidises phenolic substrates using oxygen as the electron acceptor

C:Superfamily: laccase

C:Keywords: copper; oxidoreductase

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-557/Product: laccase #status predicted <MAT>

F:80,461/Binding site: copper (His) (type 2) #status predicted

F:82,125,127,463,520,522/Binding site: 2Cu-O cluster (His) (copper type 3) #status predicted

F:458,521,526,533/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted

Query Match

Best Local Similarity 50.6%; Score 40; DB 2; Length 557;

Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 PPTCSQANSGRIST 12

Db 366 PPTCSQANSGRIST 376

RESULT 12

S32799

Hypothetical protein 1 - Xanthomonas sp. transposon Tn5053 (fragment)

C:Species: Xanthomonas sp.

C:Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 11-Jan-2000

C:Accession: S32799

R:Khododil, G.Y.; Yurleva, O.V.; Lomovskaya, O.L.; Gorlenko, Z.M.; Mindlin, S.Z.; Nikifor

J. Mol. Biol. 230, 1103-1107, 1993

A:Title: Tn5053, a mercury resistance transposon with integrin's ends.

A:Reference number: S32795; MUID:93253772; PMID:8387603

A:Accession: S32799

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-203 <KHO>

A:Cross-references: EMBL:L03735; NID:g154911; PIDN:AAA91612.1; PID:g154912

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1992

C:Genetics:

A:Mobile element: transposon Tn5053

C:Superfamily: Klebsiella transposase

Query Match

Best Local Similarity 49.4%; Score 39; DB 2; Length 203;

Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 CSOANSGRIST 14
 ||:|||||:
 Db 171 CAHASSGRVAT 181

RESULT 13

KRSML2

Keratin type II, microfibrillar - sheep

N/Alternate names: intermediate filament protein KII-9; low-sulfur keratin

C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C/Date: 15-Nov-1984 #sequence_revision 15-May-1998 #text_change 10-Dec-1999

C/Accession: 146409; A02954; S22025

R/Powell, B.; Crocker, L.; Rogers, G.

Development 114, 417-433, 1992

A/Title: Hair follicle differentiation: expression, structure and evolutionary conservat

A/Reference number: 146409; MUID:92274852; PMID:1375545

A/Accession: 146409

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-508 <PO2>

A/Cross-references: EMBL:X62509; NID:g1307; PIDN:CAA44368.1; PID:g1308

R/Crewther, W.G.; Inglis, A.S.; McKern, N.M.

Biochem. J. 173, 365-371, 1978

A/Title: Amino acid sequences of alpha-helical segments from S-carboxymethylkerateine-A.

A/Reference number: A02954; MUID:79020757; PMID:581264

A/Accession: A02954

A/Molecule type: protein

A/Residues: 142-190,192-251 <CR>

A/Note: 51-Lys and 52-Lys were also found

C/Genetics:

A/Gene: KII-9

A/Intons: 123/3; 193/3; 214/3; 246/3; 301/3; 343/3; 417/2; 428/1

C/Superfamily: Cytoskeletal keratin

C/Keywords: coiled coil; cytoskeleton; hair; intermediate filament

F/126-413/Domain: rod #status predicted <ROD>

F/126-253/Region: coiled coil 1 #status predicted

F/271-418/Region: coiled coil 2 #status predicted

Query Match 49.4%; Score 39; DB 1; Length 508;
 Best Local Similarity 54.5%; Pred. No. 1.1e+02;
 Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

RESULT 14

E95889

Probable dehydrogenase protein [imported] - Sinorhizobium meliloti (strain 1021) magapla

C/Species: Sinorhizobium meliloti

C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C/Accession: E95889

R/Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A/Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo

A/Reference number: A95842; MUID:21396508; PMID:11481431

A/Accession: E95889

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-749 <KUR>

A/Cross-references: GB:U59185; PIDN:CA48781.1; PID:g15140254; GSPDB:GN00167

A/Experimental source: strain 1021, megaplasmid pSymB

R/Gailbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hublier,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaulic, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.

A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A/Reference number: A96039; MUID:21368234; PMID:11474104

A/Contents: annotation

C/Genetics:
 A/Gene: Smb20395
 A/Genome: plasmid

Query Match 49.4%; Score 39; DB 2; Length 749;
 Best Local Similarity 50.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 PTCOANSGRISTL 15
 |||||:
 Db 49 PVSCTIANGRIITTI 62

RESULT 15

JC2129

protein kinase PKN (EC 2.7.1.-) - human

N/Alternate names: protein-kinase-C-related kinase PKR1

C/Species: Homo sapiens (man)

C/Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 17-Nov-2000

C/Accession: JC2129; S67526; S51020; I53327

R/Mukai, H.; Ono, Y.

Biochem. Biophys. Res. Commun. 199, 897-904, 1994

A/Title: A novel protein kinase with leucine zipper-like sequences: Its catalytic domain

A/Reference number: JC2129; MUID:94183274; PMID:8135837

A/Accession: JC2129

A/Molecule type: mRNA

A/Residues: 1-942 <MK>

A/Cross-references: GB:D26181; NID:g473712; PIDN:BAA05169.1; PID:g825505

R/Palmer, R.H.; Ridden, U.; Parker, P.O.

Eur. J. Biochem. 227, 344-351, 1995

A/Title: Cloning and expression patterns of two members of a novel protein-kinase-C-rela

A/Reference number: I53327; MUID:95154310; PMID:7851406

A/Accession: S67526

A/Molecule type: mRNA

A/Residues: 1-190, 'D', 192-942 <PAL>

A/Cross-references: EMBL:S75546; NID:g914097; PIDN:AAB33345.1; PID:g914098

A/Experimental source: fetal brain

R/Palmer, R.H.; Ridden, U.; Parker, P.O.

FEBS Lett. 356, 5-8, 1994

A/Title: Identification of multiple, novel, protein kinase C-related gene products.

A/Reference number: S51020; MUID:95080426; PMID:7988719

A/Accession: S51020

A/Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra

A/Molecule type: mRNA

A/Residues: 700-799, 'A' <PA2>

C/Comment: This protein has a protein kinase domain related to protein kinase C.

C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolc

C/Keywords: ATP; leucine zipper; phosphotransferase; protein kinase

F/39-66/Region: basic

F/70-287/Region: leucine zipper motif

F/613-874/Domain: protein kinase ATP-binding motif

F/621-629/Region: protein kinase ATP-binding motif

F/644/Active site: Lys #status predicted

Query Match 49.4%; Score 39; DB 2; Length 942;
 Best Local Similarity 50.0%; Pred. No. 1.9e+02;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 PTCOANSGRISTL 15
 |||||:
 Db 118 PTCSATNLSRVAGL 131

Search completed: November 26, 2003, 12:36:04
 Job time : 14.9157 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:25:10 ; Search time 32.7108 Seconds

(without alignments)
118.334 Million cell updates/sec

Title: US-09-230-111C-14

Sequence: 1 PPTCQANSGRSTL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_RVIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	62.0	336	11 Q8VE51	Q8VE51 mus musculus
2	49	62.0	336	11 Q8CI12	Q8CI12 mus musculus
3	47	59.5	336	11 Q62834	Q62834 rattus norv
4	44	55.7	409	10 P93368	P93368 nicotiana t
5	44	55.7	1866	10 Q85ER3	Q85ER3 oryza sativ
6	43	54.4	465	10 Q8VZY0	Q8VZY0 oryza sativ
7	43	54.4	496	10 Q8GS39	Q8GS39 oryza sativ
8	43	54.4	591	13 Q8QFW6	Q8QFW6 raja eglant
9	43	54.4	699	16 P95147	P95147 mycobacteri
10	42	53.2	504	11 Q8BIF4	Q8BIF4 mus musculu
11	42	53.2	1203	5 Q9NSK0	Q9NSK0 caenorhabdi
12	42	53.2	1257	11 Q92301	Q92301 rattus norv
13	42	53.2	1844	5 Q22579	Q22579 caenorhabdi
14	41	51.9	133	4 Q9Y3Y1	Q9Y3Y1 homo sapien
15	41	51.9	277	5 Q44870	Q44870 caenorhabdi
16	41	51.9	1263	13 Q8J1Y6	Q8J1Y6 brachydanio

17	40	50.6	304	16 Q86604	Q86604 streptomyce
18	40	50.6	419	10 Q9SVT7	Q9SVT7 arabidopsis
19	40	50.6	557	10 P93366	P93366 nicotiana t
20	40	50.6	624	10 Q855X8	Q855X8 oryza sativ
21	40	50.6	1324	5 Q9W1L2	Q9W1L2 drosophila
22	39	49.4	176	16 Q8G7T5	Q8G7T5 bifidobacte
23	39	49.4	222	11 Q9D857	Q9D857 mus musculu
24	39	49.4	245	5 Q18668	Q18668 drosophila
25	39	49.4	253	5 Q9TVF2	Q9TVF2 drosophila
26	39	49.4	255	5 Q9U4N8	Q9U4N8 drosophila
27	39	49.4	255	5 Q9U4N8	Q9U4N8 drosophila
28	39	49.4	275	12 Q8VZ01	Q8VZ01 camelipox vi
29	39	49.4	279	11 Q9CZD0	Q9CZD0 mus musculu
30	39	49.4	340	4 Q8H8T8	Q8H8T8 homo sapien
31	39	49.4	348	12 Q99A00	Q99A00 carnation m
32	39	49.4	508	6 Q28582	Q28582 ovis aries
33	39	49.4	542	10 Q8LSW7	Q8LSW7 helianthus
34	39	49.4	547	4 Q81V81	Q81V81 homo sapien
35	39	49.4	749	16 Q92W88	Q92W88 rhizobium m
36	39	49.4	940	10 Q9SD55	Q9SD55 oryza sativ
37	39	49.4	942	4 Q81UV5	Q81UV5 homo sapien
38	39	49.4	1559	13 Q73697	Q73697 tugu rubrip
39	38.5	48.7	426	10 Q8H407	Q8H407 oryza sativ
40	38	48.1	84	17 Q97XN0	Q97XN0 sulfobobus
41	38	48.1	110	5 Q9XX22	Q9XX22 caenorhabdi
42	38	48.1	122	12 Q993U0	Q993U0 porcine rep
43	38	48.1	123	12 Q9YX89	Q9YX89 porcine rep
44	38	48.1	123	12 Q90051	Q90051 porcine rep
45	38	48.1	123	12 Q70657	Q70657 porcine rep

ALIGNMENTS

RESULT 1

Q8VE51 PRELIMINARY; PRT; 336 AA.

AC Q8VE51
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to D123 gene product (Hypothetical protein) (Expressed
in sequence AA959893) (D123 homolog).
GN AA959893.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RA "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RT Nature 420:563-573 (2002).
DR EMBL, BC019753; AAH19753.1;
DR EMBL, BC022972; AAH22972.1;
DR EMBL, BC024787; AAH24787.1; -.

DR EMBL; AK090122; BAC41103.1; -
 DR MGD; MGI:2138811; AA959893.
 KM Hypothetical protein
 SQ SEQUENCE 336 AA; 38816 MW; 1E59461A424B7A76 CRC64;

Query Match
 Best Local Similarity 62.0%; Score 49; DB 11; Length 336;
 Pred. No. 0.96;
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPTCSQANSNG 10
 DB 52 PPTCSQSDSG 61

RESULT 2
 ID Q8C112 PRELIMINARY; PRT; 336 AA.
 AC Q8C112;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to expressed sequence AA959893.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCB1_TaxID=10090;
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N;
 RA Strusberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC023821; AAH23821.1; -
 SQ SEQUENCE 336 AA; 38804 MW; 1E48B616B2567A76 CRC64;

Query Match
 Best Local Similarity 62.0%; Score 49; DB 11; Length 336;
 Pred. No. 0.96;
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPTCSQANSNG 10
 DB 52 PPTCSQSDSG 61

RESULT 3
 ID Q62834 PRELIMINARY; PRT; 336 AA.
 AC Q62834;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE D123.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN NCB1_TaxID=10116;
 RP SEQUENCE FROM N.A.
 RC STRAIN=Fischer;
 RX MEDLINE=96177290; PubMed=8601400;
 RA Okuda A.; Kimura G.;
 RT "An amino acid change in novel protein D123 is responsible for
 temperature-sensitive G1-phase arrest in a mutant of rat fibroblast
 line 3Y1.";
 RL Exp. Cell Res. 223:242-249(1996).
 DR EMBL; U34843; AAB60521.1; -
 FT VARIANT 109 109 A -> V.
 SQ SEQUENCE 336 AA; 38812 MW; 6E1D1379A2612934 CRC64;

Query Match
 Best Local Similarity 59.5%; Score 47; DB 11; Length 336;
 Pred. No. 2.2;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPTCSQANSNG 10

DB 52 PPTCSQSDSG 61

RESULT 4
 ID P93368 PRELIMINARY; PRT; 409 AA.
 AC P93368;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Diphenol oxidase (EC 1.10.3.2) (Fragment).
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.
 RN NCB1_TaxID=4097;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97080572; PubMed=8921917;
 RA Kiefer-Meyer M.C.; Gomord V.; O'Connell A.; Halpin C.; Faye L.;
 RT "Cloning and sequence analysis of laccase-encoding cDNA clones from
 tobacco.";
 RL Gene 178:205-207(1996).
 DR EMBL; U45243; AAC49538.1; -
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR002355; Multicu_oxidase2.
 DR Pfam: PF00394; Cu-oxidase; 2.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 1.
 DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; 1.
 KW Oxidoreductase.
 FT NON TER 1
 SQ SEQUENCE 409 AA; 44056 MW; 6254ABAFB509D4AD CRC64;

Query Match
 Best Local Similarity 55.7%; Score 44; DB 10; Length 409;
 Pred. No. 9.8;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PPTCSQANSGR 12
 DB 218 PPTCSQANSRV 228

RESULT 5
 ID Q85SR3 PRELIMINARY; PRT; 1866 AA.
 AC Q85SR3;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Putative copia-type pol polyprotein.
 GN OJ136B01.16.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoideae; Oryza.
 RN NCB1_TaxID=4530;
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nipponbare;
 RX McCombe W.R.; de la Bastide M.; Spiegel L.; Preson R.; Kirchoff K.;
 RA Kuit K.; Nascimento L.; Zuberavert T.; Balija V.; Bell M.; Baker J.;
 RA Santos L.; Miller B.; Katzenberger F.; Muller S.; King L.; Yang C.;
 RA O'Shaughnessy A.; Palmer L.; Dedhia N.;
 RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone
 OJ136B01, from chromosome 10, complete sequence.";
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC108883; AAM08642.1; -

Query Match
 Best Local Similarity 59.5%; Score 47; DB 11; Length 336;
 Pred. No. 2.2;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DR SMART, SM00343; ZNF_C2HC; 1.
 DR PROSITE; PS50158; 2F_CCHC; 1.
 KW POLYPEPTIDE
 SQ SEQUENCE 1866 AA; 212916 MW; F66634B07CA9467E CRC64;

Query Match 55.7%; Score 44; DB 10; Length 1866;
 Best Local Similarity 60.0%; Pred. No. 47;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PPTCSQANSGRISTL 15
 |||:|||||:
 Db 842 PPTSTQANDPSISL 856

RESULT 6

Q8VZYO PRELIMINARY; PRT; 465 AA.
 AC Q8VZYO;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE C13 cysteine proteinase.
 GN NP1.
 OS Oryza sativa (indica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriactoidae; Oryzaceae; Oryza.
 NCBI_TaxID=39946;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bi X.-Z., Khush G.S., Bennett J.;
 RT "Cloning and expression of barley nucellin, C13 endopeptidase NP1
 precursor from rice ovule."
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY062178; AAL40390.1; -
 DR Gramene; Q8VZYO; -
 DR InterPro; IPR001096; Legumain.
 DR Pfam; PF01650; Peptidase_C13; 1.
 DR PRINTS; PR00776; HEMOGLOBINASE.
 SQ SEQUENCE 465 AA; 51906 MW; 8B13C140753DD33 CRC64;

Query Match 54.4%; Score 43; DB 10; Length 465;
 Best Local Similarity 69.2%; Pred. No. 17;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 TCSQANSGRISTL 15
 |||:|||||:
 Db 447 TCSHNSGRWSSL 459

RESULT 7

Q8G639 PRELIMINARY; PRT; 496 AA.
 AC Q8G639;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Asparaginyl endopeptidase REP-2.
 GN REP2.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriactoidae; Oryzaceae; Oryza.
 NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Kato H., Sutoh K., Minamikawa T.;
 RT "Identification, cDNA Cloning and Possible Roles of Seed Specific Rice
 Asparaginyl Endopeptidase, REP-2."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB081464; BAC41386.1; -
 DR EMBL; AB081465; BAC41387.1; -

SQ SEQUENCE 496 AA; 54848 MW; 7E583C3E8AB886D0D CRC64;

Query Match 54.4%; Score 43; DB 10; Length 496;
 Best Local Similarity 69.2%; Pred. No. 18;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 TCSQANSGRISTL 15
 |||:|||||:
 Db 478 TCSHNSGRWSSL 490

RESULT 8

Q8QFW6 PRELIMINARY; PRT; 591 AA.
 AC Q8QFW6;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE BBL.
 OS Raja eglaneria (cleamose skate).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Squalia; Hypnosqualia; Pristiogaster; Batoidae;
 OC Rajiformes; Rajidae; Raja.
 NCBI_TaxID=33514;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Anderson M.K., Miracle A.L., Sun X., Luer C.A., Walsh C.J.,
 RA Litman G.W., Rothenberg E.V.;
 RT "Early vertebrate origins of lymphocyte developmental regulatory
 pathways."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY079197; AAL86576.1; -
 DR InterPro; IPR001092; HLH_basic.
 DR InterPro; IPR002909; IPT_TIG.
 DR InterPro; IPR003523; TF_COE.
 DR Pfam; PF01833; TIG; 1.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00429; IPT; 1.
 DR PROSITE; PS01345; COE; 1.
 SQ SEQUENCE 591 AA; 64413 MW; 826411F6B979E2B6 CRC64;

Query Match 54.4%; Score 43; DB 13; Length 591;
 Best Local Similarity 40.0%; Pred. No. 22;
 Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 PPTCSQANSGRISTL 15
 |||:|||||:
 Db 569 PPTCSANANSLQAI 583

RESULT 9

P95147 PRELIMINARY; PRT; 699 AA.
 AC P95147;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Hypothetical protein Rvi1868.
 GN RVI1868 OR MTCY359.05C OR MTI917.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Broch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gae S., Barry C.E. III, Tekala F.,
 RA Davies R., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Badcock K., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

```

GN H43E16.1
OC Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;
OC Rhabdittidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Miller N., Carter T.;
RT "The sequence of C. elegans cosmid H43E16,";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Waterston R.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006669; AAF39909.1; -
DR WormPeP; H43E16.1; CE19949.
KW Hypothetical protein.
SQ SEQUENCE 1203 AA; 120104 MW; E04CADC2BA74713B CRC64;

QY 2 PTCQANSGRISTL 15
DB 1017 PTCASPYSGKIATV 1030

53.2%; Score 42; DB 5; Length 1203;
Best local similarity 50.0%; Pred. No. 70;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

RESULT 12
092301
ID 092301 PRELIMINARY; PRT; 1257 AA.
AC 092301;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE RPER2.
GN RPER2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=98438460; PubMed=9765215;
RA Sakamoto K., Nagase T., Fukui H., Horikawa K., Okada T., Tanaka H.,
RA Sato K., Miyake Y., Ohara O., Kato K., Ishida N.;
RT "Multitissue circadian expression of rat period homologue (rPer2) mRNA
RT is governed by the mammalian circadian clock, the suprachiasmatic
RT nucleus in the brain.";
RL J. Biol. Chem. 273:27039-27042 (1998).
DR EMBL; AB016532; BAA34187.1; -.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR00014; PAS_domain.
DR SMART; SMO0086; PAC; 1.
DR SMART; SMO0091; PAS; 1.
DR PROSITE; PS0112; PAS; 1.
SQ SEQUENCE 1257 AA; 136027 MW; A772B3C45E63CED CRC64;

Query Match 53.2%; Score 42; DB 11; Length 1257;
Best local similarity 55.6%; Pred. No. 73;
Matches 10; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

```

QY 1 PTCSEQ---ANSGRIST 14
 |||||: |||||
 DB 1029 PTCSEPTSDONDAIST 1046

RESULT 13

Q22579 PRELIMINARY; PRT; 1844 AA.
 AC Q22579;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical 187.1 kDa protein.
 GN T19D12.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodertinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX Favello A.;
 RT "The sequence of C. elegans cosmid T19D12.";
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U41263; AAC24428.2; -.
 DR WormPep; T19D12.1; CE28677.
 KW Hypothetical protein.
 SQ SEQUENCE 1844 AA; 187133 MW; 6151A72777517F5 CRC64;

Query Match 53.2%; Score 42; DB 5; Length 1844;
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 PTCSEQSGRISTL 15
 |||||: |||||
 DB 1659 PTCASPYSGKATV 1672

RESULT 14
 Q93Y1 PRELIMINARY; PRT; 133 AA.
 AC Q93Y1;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Hypothetical protein (Fragment).
 GN DKFZPS86E1422.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Uterus;
 RA Koehler K., Beyer A., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL050170; CAB43306.1; -.
 KW Hypothetical protein.
 FT NON_TER 1

SQ SEQUENCE 133 AA; 14307 MW; F098777EB4707C61 CRC64;

Query Match 51.9%; Score 41; DB 4; Length 133;
 Best Local Similarity 63.6%; Pred. No. 11;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PTCSEQSGRI 12
 |||||: |||||
 DB 115 PACSQSGGAV 125

RESULT 15

O44870 PRELIMINARY; PRT; 277 AA.
 AC O44870;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical 31.6 kDa protein.
 GN K05F6.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodertinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Du Z., Goela D.;
 RT "The sequence of C. elegans cosmid K05F6.";
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF040653; AAB35023.1; -.
 DR WormPep; K05F6.10; CE17168.
 KW Hypothetical protein.
 SQ SEQUENCE 277 AA; 31566 MW; AB6052724B484C2C CRC64;

Query Match 51.9%; Score 41; DB 5; Length 277;
 Best Local Similarity 57.1%; Pred. No. 23;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 PTCSEQSGRIST 14
 |||||: |||||
 DB 51 PPTAKANFGREGT 64

Search completed: November 26, 2003, 12:34:37
 Job time : 34.7108 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 26, 2003, 12:26:20 ; Search time 27.1084 Seconds
(Without alignments)
102.059 Million cell updates/sec

Title: US-09-230-111C-15

Perfect score: 71
Sequence: 1 IDLASEFLFLSNSFL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 673684 seqs, 184443283 residues

Total number of hits satisfying chosen parameters: 673684

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71	100.0	15	8 US-08-681-219-17	Sequence 17, Appl
2	71	100.0	15	12 US-10-092-113-15	Sequence 15, Appl
3	71	100.0	15	12 US-09-230-111C-15	Sequence 15, Appl
4	37	52.1	43	11 US-09-983-802-264	Sequence 264, Appl
5	37	52.1	80	15 US-10-000-256A-178	Sequence 178, Appl
6	37	52.1	245	10 US-09-954-737-9	Sequence 9, Appl
7	36	50.7	346	9 US-09-741-669-421	Sequence 421, Appl
8	36	50.7	346	9 US-09-815-242-10206	Sequence 10206, A
9	36	50.7	468	15 US-10-156-761-9362	Sequence 9362, Ap
10	36	50.7	498	12 US-09-766-511B-26	Sequence 26, Appl
11	36	50.7	498	12 US-10-188-495-51	Sequence 51, Appl
12	36	50.7	498	15 US-10-189-123-51	Sequence 51, Appl
13	36	50.7	552	9 US-09-764-898-183	Sequence 183, Appl
14	36	50.7	553	9 US-09-764-898-163	Sequence 163, Appl
15	36	50.7	591	12 US-09-766-511B-25	Sequence 25, Appl

16	36	50.7	591	12	US-10-188-495-50	Sequence 50, Appl
17	36	50.7	591	15	US-10-189-123-50	Sequence 50, Appl
18	36	50.7	622	11	US-09-764-881-119	Sequence 119, Appl
19	36	50.7	622	12	US-10-017-161-1176	Sequence 1176, Ap
20	36	50.7	622	12	US-10-137-870-180	Sequence 180, Appl
21	36	50.7	622	12	US-10-140-018-180	Sequence 180, Appl
22	36	50.7	622	12	US-10-140-021-180	Sequence 180, Appl
23	36	50.7	622	12	US-10-140-274-180	Sequence 180, Appl
24	36	50.7	622	12	US-10-140-471-180	Sequence 180, Appl
25	36	50.7	622	12	US-10-140-807-180	Sequence 180, Appl
26	36	50.7	622	12	US-10-140-922-180	Sequence 180, Appl
27	36	50.7	622	12	US-10-140-924-180	Sequence 180, Appl
28	36	50.7	622	12	US-10-140-926-180	Sequence 180, Appl
29	36	50.7	622	12	US-10-141-698-180	Sequence 180, Appl
30	36	50.7	622	12	US-10-141-702-180	Sequence 180, Appl
31	36	50.7	622	12	US-10-141-704-180	Sequence 180, Appl
32	36	50.7	622	12	US-10-142-432-180	Sequence 180, Appl
33	36	50.7	622	12	US-10-142-433-180	Sequence 180, Appl
34	36	50.7	622	12	US-10-142-767-180	Sequence 180, Appl
35	36	50.7	622	12	US-10-143-033-180	Sequence 180, Appl
36	36	50.7	622	12	US-10-143-994-180	Sequence 180, Appl
37	36	50.7	622	12	US-10-145-628-180	Sequence 180, Appl
38	36	50.7	622	12	US-10-145-631-180	Sequence 180, Appl
39	36	50.7	622	12	US-10-145-633-180	Sequence 180, Appl
40	36	50.7	622	12	US-10-145-746-180	Sequence 180, Appl
41	36	50.7	622	12	US-10-145-748-180	Sequence 180, Appl
42	36	50.7	622	12	US-10-145-823-180	Sequence 180, Appl
43	36	50.7	622	12	US-10-145-826-180	Sequence 180, Appl
44	36	50.7	622	12	US-10-145-870-180	Sequence 180, Appl
45	36	50.7	622	12	US-10-145-876-180	Sequence 180, Appl

ALIGNMENTS

RESULT 1
US-08-681-219-17
Sequence 17, Application US/08681219
Publication No. US20020058607A1
GENERAL INFORMATION:
APPLICANT: Takaaki Sato and Junn Yanagisawa
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GUGP
TITLE OF INVENTION: (PDE/DHR) DOMAIN AND USBS THEREOF
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESS: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/681,219
FILING DATE: 22-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/48962/JPM/JKM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ. ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-681-219-17

Query Match 100.0%; Score 71; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IDLASEFLFLSNSFL 15
DB 1 IDLASEFLFLSNSFL 15

RESULT 2
US-10-092-138-15
Sequence 15; Application US/10092138
Publication No. US2003010723A1
GENERAL INFORMATION:

APPLICANT: Sato, Taka-Aki
TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
TITLE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION
FILE REFERENCE: 65823/JPW/PT
CURRENT APPLICATION NUMBER: US/10/092,138
CURRENT FILING DATE: 2002-09-06
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial
US-10-092-138-15
OTHER INFORMATION: Sequence:source:synthesized

Query Match 100.0%; Score 71; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IDLASEFLFLSNSFL 15
DB 1 IDLASEFLFLSNSFL 15

RESULT 3
US-09-230-111C-15
Sequence 15; Application US/09230111C
Publication No. US2003020341A1
GENERAL INFORMATION:

APPLICANT: Sato, Taka-Aki
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN
TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF (PDZ/DHR)
TITLE OF INVENTION: DOMAIN AND USES THEREOF
FILE REFERENCE: 48962-A-PCT-US
CURRENT APPLICATION NUMBER: US/09/230,111C
CURRENT FILING DATE: 1999-05-17
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial
US-09-230-111C-15
OTHER INFORMATION: Sequence:source:synthesized

Query Match 100.0%; Score 71; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IDLASEFLFLSNSFL 15

DB 1 IDLASEFLFLSNSFL 15

RESULT 4
US-09-983-802-264
Sequence 264; Application US/09983802
Publication No. US20030022185A1
GENERAL INFORMATION:

APPLICANT: Fischer et al.
TITLE OF INVENTION: 123 Human Secreted Proteins
FILE REFERENCE: P2010P1
CURRENT APPLICATION NUMBER: US/09/983,802
CURRENT FILING DATE: 2001-10-25
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/227,357
PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US98/13664
PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,926
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,793
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,925
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,929
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,803
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,732
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,931
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,932
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,916
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,930
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,918
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,920
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,733
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,795
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,919
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,928
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,722
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,723
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,948
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,949
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,953
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,950
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,947
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,964
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/056,360
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,684
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,984
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,954

```

; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,785
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-13
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,664
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,660
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,661
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 264
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (43)
; OTHER INFORMATION: Xaa equals stop translation
US-09-983-802-264

```

```

Query Match      52.1%; Score 37; DB 11; Length 43;
Best Local Similarity 60.0%; Pred. No. 28;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      1 IDLASEPFLSNSFL 15
       : ||| ||| |||
Db      9 LSLSLFLFLPPSFL 23

```

```

RESULT 5
US-10-000-256A-178
; Sequence 178, Application US/10000256A
; Publication No. US2003003983A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Harvey
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0259
; CURRENT APPLICATION NUMBER: US/10/000,256A
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 60/244,782
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 178
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-000-256A-178

```

```

Query Match      52.1%; Score 37; DB 15; Length 80;
Best Local Similarity 61.5%; Pred. No. 53;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      3 LASEFLFLSNSFL 15
       : ||| ||| |||
Db      27 LASHFLSLSSSYL 39

```

```

RESULT 6
US-09-954-737-9
; Sequence 9, Application US/09954737
; Patent No. US2002010860A1
; GENERAL INFORMATION:
; APPLICANT: Byron, Sierd
; APPLICANT: Jongbloed, Jan D.H.
; APPLICANT: Mueller, Joerg P.
; APPLICANT: Van Dijk, Jan M.
; TITLE OF INVENTION: Twin-Arginine Translocation in Bacillus
; FILE REFERENCE: GC634-2

```

```

; CURRENT APPLICATION NUMBER: US/09/954,737
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: US 60/233,610
; PRIOR FILING DATE: 2000-09-18
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-954-737-9

```

```

Query Match      52.1%; Score 37; DB 10; Length 245;
Best Local Similarity 53.3%; Pred. No. 16e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      1 IDLASEPFLSNSFL 15
       : ||| ||| |||
Db      21 VTLAAFLFLITRAFL 35

```

```

RESULT 7
US-09-741-669-421
; Sequence 421, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Foreyth, R. Allyn
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyckind, Judith W.
; TITLE OF INVENTION: Genes identified as required for
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 421
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-741-669-421

```

```

Query Match      50.7%; Score 36; DB 9; Length 346;
Best Local Similarity 57.1%; Pred. No. 3.4e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 IDLASEPFLSNSFL 14
       : ||| ||| |||
Db      191 IDISSEKLAAXSF 204

```

```

RESULT 8
US-09-815-242-10206
; Sequence 10206, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haseelbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyckind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21

```

```
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 10206
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10206
```

```
Query Match          50.7%; Score 36; DB 9; Length 346;
Best Local Similarity 57.1%; Pred. No. 3.4e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 IDLASEFLSNSF 14
||:|||||:|
Db 191 IDISEKALAKSF 204
```

```
RESULT 9
US-10-156-761-9362
; Sequence 9362, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO: 9362
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Streptomyces avermectinis
US-10-156-761-9362
```

```
Query Match          50.7%; Score 36; DB 15; Length 468;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 2 DLASEFLSNSF 15
||:|||||:|
Db 359 DVKSEWLFPAFL 372
```

```
RESULT 10
US-09-766-511B-26
; Sequence 26, Application US/09766511B
; Publication No. US20030170621A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
```

```
; APPLICANT: BARNES, Thomas S
; APPLICANT: KIRST, Susan J
; APPLICANT: MYERS, Paul S
; APPLICANT: WRIGHTON, Nicholas
; APPLICANT: GOODEARL, Andrew
; APPLICANT: HOLTZMAN, Douglas A
; APPLICANT: KHODADUST, Mehran M
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING PROGNOSTIC, DIAGNOSTIC, PREV
; TITLE OF INVENTION: THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-65
; CURRENT APPLICATION NUMBER: US/09/766,511B
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/345,680
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 26
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-766-511B-26
```

```
Query Match          50.7%; Score 36; DB 12; Length 498;
Best Local Similarity 66.7%; Pred. No. 4.9e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 7 FFLSNSF 15
|||:||||:|
Db 81 FFLNNF 89
```

```
RESULT 11
US-10-188-495-51
; Sequence 51, Application US/10188495
; Publication No. US2003017573A1
; GENERAL INFORMATION:
; APPLICANT: KIRST, Susan J.
; APPLICANT: HOLTZMAN, Douglas A.
; APPLICANT: FRASER, Christopher C.
; APPLICANT: SHARP, John D.
; APPLICANT: BARNES, Thomas S.
; TITLE OF INVENTION: POLYPEPTIDES HAVING DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTH
; FILE REFERENCE: 10147-11U2
; CURRENT APPLICATION NUMBER: US/10/188,495
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 51
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-495-51
```

```
Query Match          50.7%; Score 36; DB 12; Length 498;
Best Local Similarity 66.7%; Pred. No. 4.9e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```


Qy 7 FFLSNSFL 15
|||:|:|:
Db 81 FFLNNFT 89

RESULT 12

US-10-189-123-51
; Sequence 51, Application US/10189123
; Publication No. US20030082586A1
; GENERAL INFORMATION:
; APPLICANT: KIRST, Susan J.
; APPLICANT: HOLTZMAN, Douglas A.
; APPLICANT: FRASER, Christopher C.
; APPLICANT: SHARP, John D.
; APPLICANT: BARNES, Thomas S.
; TITLE OF INVENTION: ANTIBODIES HAVING DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE REFERENCE: 10147-11U3
; CURRENT APPLICATION NUMBER: US/10/189,123
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 09/342,364
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 51
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-189-123-51

Query Match 50.7%; Score 36; DB 15; Length 498;
Best Local Similarity 66.7%; Pred. No. 4.9e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 7 FFLSNSFL 15
|||:|:|:
Db 81 FFLNNFT 89

RESULT 13

US-09-764-898-183
; Sequence 183, Application US/09764898
; Patent No. US20020090673A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P201
; CURRENT APPLICATION NUMBER: US/09/764,898
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 183
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-898-183

Query Match 50.7%; Score 36; DB 9; Length 552;
Best Local Similarity 66.7%; Pred. No. 5.5e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IDLASEFLPLSN 12
:|:|:|:|:
Db 130 IDLACEFLCLLN 141

RESULT 14
US-09-764-898-163
; Sequence 163, Application US/09764898
; Patent No. US20020090673A1
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P201
; CURRENT APPLICATION NUMBER: US/09/764,898
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 163
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (393)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-898-163

Query Match 50.7%; Score 36; DB 9; Length 553;
Best Local Similarity 66.7%; Pred. No. 5.5e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IDLASEFLPLSN 12
:|:|:|:|:
Db 131 IDLACEFLCLLN 142

RESULT 15
US-09-766-511B-25
; Sequence 25, Application US/09766511B
; Publication No. US20030170621A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; APPLICANT: BARNES, Thomas S
; APPLICANT: KIRST, Susan J
; APPLICANT: MYERS, Paul S
; APPLICANT: WRIGHTON, Nicholas
; APPLICANT: GOODEARL, Andrew
; APPLICANT: HOLTZMAN, Douglas A
; APPLICANT: KHODADOUST, Mehrian M
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING PROGNOSTIC, DIAGNOSTIC, PRE
; TITLE OF INVENTION: THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-65
; CURRENT APPLICATION NUMBER: US/09/766,511B
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/345,680
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 25
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-766-511B-25

Query Match 50.7%; Score 36; DB 12; Length 591;
Best Local Similarity 66.7%; Pred. No. 5.9e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy	7	FLFISNSFL	15
Db	81	FLFISNSFL	89

Search completed: November 26, 2003, 12:38:40
Job time : 28.1084 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:21:29 ; Search time 41.5663 Seconds
(without alignments)
57.280 Million cell updates/sec

Title: US-09-230-111c-15
Perfect score: 71
Sequence: 1 IDLASERFLNSRL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	71	100.0	15	19	AAW50174
2	45	63.4	682	21	AAW52645
3	45	63.4	682	23	ABW92235
4	45	63.4	725	21	AAW39851
5	42	59.2	60	23	ABW03996
6	42	59.2	106	20	AAW55732
7	42	59.2	106	20	AAW55733
8	41	57.7	465	22	ABW03158
9	40	56.3	183	22	AAW19348

10	40	56.3	280	23	ABW30033	Streptococcus poly
11	40	56.3	554	22	AAU14133	Human novel protei
12	38	53.5	53	22	AAW90853	Human Immune/haema
13	38	53.5	134	24	ABW00902	S. pneumoniae type
14	38	53.5	253	22	ABW05253	Novel human diagno
15	38	53.5	467	22	ABW63044	Drosophila melanog
16	38	53.5	527	22	ABW99962	ERA binding domain
17	38	53.5	548	23	ABW48353	Listeria monocyrog
18	37	52.1	43	20	AAW02764	Human secreted pro
19	37	52.1	80	23	ABW77072	Prostate specific
20	37	52.1	202	23	ABW91656	Herbicidally activ
21	37	52.1	245	23	AAW80095	Putative P. abyssi
22	37	52.1	309	22	ABW96216	Staphylococcus epi
23	37	52.1	901	23	ABW40506	Human immune/haema
24	37	52.1	1178	18	AAW30763	Human immune/haema
25	36	50.7	44	22	AAW85574	Novel human diagno
26	36	50.7	59	22	ABW20430	Human secreted pro
27	36	50.7	60	21	AAW45030	Antigen GX3373 enc
28	36	50.7	76	11	AAW05218	Human immune/haema
29	36	50.7	84	22	AAW84660	Human polypeptide
30	36	50.7	133	22	AAW04230	N. gonorrhoeae ami
31	36	50.7	164	24	ABW77460	Human G-Protein Co
32	36	50.7	189	22	AAW27513	Human G-Protein Co
33	36	50.7	213	22	AAW27534	E. coli cellular p
34	36	50.7	346	22	AAW34613	Escherichia coli p
35	36	50.7	346	22	AAW98373	Putative UDP-N-ace
36	36	50.7	418	22	AAW96180	S cerevisiae apopt
37	36	50.7	498	22	AAW61228	Human novel secret
38	36	50.7	517	22	AAW70715	Human novel secret
39	36	50.7	552	22	AAW16942	Human novel secret
40	36	50.7	553	22	AAW16922	Mature human TANGO
41	36	50.7	591	22	AAW61227	Mature human neuro
42	36	50.7	597	22	AAW06798	Human neuronal gui
43	36	50.7	622	22	AAW06789	Human PRQ4337 poly
44	36	50.7	622	22	AAW12261	Human TANGO 325 pr
45	36	50.7	622	22	AAW61225	

ALIGNMENTS

RESULT 1

AAW50174	standard; peptide; 15 AA.
ID	AAW50174
AC	AAW50174;
XX	
DT	16-JUL-1998 (first entry)
DE	Signal-transducing protein carboxy-terminal peptide.
XX	
XX	Inhibition; specific binding; signal-transducing protein;
KW	cytoplasmic protein; proliferation; cancer cell; apoptosis;
KW	virally infected cell.
OS	Synthetic.
PN	W09805347-A1.
XX	
PD	12-FEB-1998.
PF	18-JUL-1997; 97WO-US12677.
XX	
PR	22-JUL-1996; 96US-0681219.
XX	
PA	(UYCO) UNIV COLUMBIA NEW YORK.
XX	
PI	Sato T, Yanagisawa J;
DR	WPI; 1998-145347/13.
XX	
PT	Inhibition of signal transduction - by inhibiting binding between a
PT	signal-transducing protein and a cytoplasmic protein, for treating

PT e.g. cancer or viral infection
XX
PS Claim 16; Page 64; 108pp; English.
XX
XX A novel composition is capable of inhibiting specific binding
CC between a signal-transducing protein (STR) having the
CC carboxy-terminal sequence (Ser/Thr)-Xaa-(Val/Ile/Leu), where Xaa =
CC any amino acid (e.g. the present peptide), and a cytoplasmic
CC protein (CP) containing the sequence AAW50162 or AAW50163.
CC The composition can be used to inhibit the proliferation of cancer
CC or virally infected cells, or induce apoptosis in cancer or virally
CC infected cells.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 71; DB 19; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IDLASEFLFNSFL 15
Db 1 IDLASEFLFNSFL 15

RESULT 2
ID AAG25645
AAG25645 strand; Protein; 682 AA.
XX
AC AAG25645;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 29791.
XX
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
XX
XX
XX 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX
XX 25-FEB-1999; 99US-0121825.
XX
XX 05-MAR-1999; 99US-0123180.
XX
XX 09-MAR-1999; 99US-0123548.
XX
XX 23-MAR-1999; 99US-0125788.
XX
XX 25-MAR-1999; 99US-0126264.
XX
XX 29-MAR-1999; 99US-0126785.
XX
XX 01-APR-1999; 99US-0127462.
XX
XX 06-APR-1999; 99US-0128234.
XX
XX 08-APR-1999; 99US-0128714.
XX
XX 16-APR-1999; 99US-0129845.
XX
XX 19-APR-1999; 99US-0130077.
XX
XX 21-APR-1999; 99US-0130449.
XX
XX 23-APR-1999; 99US-0130510.
XX
XX 28-APR-1999; 99US-0130891.
XX
XX 30-APR-1999; 99US-0131449.
XX
XX 30-APR-1999; 99US-0132407.
XX
XX 04-MAY-1999; 99US-0132484.
XX
XX 05-MAY-1999; 99US-0132485.
XX
XX 06-MAY-1999; 99US-0132486.
XX
XX 07-MAY-1999; 99US-0132863.
XX
XX 11-MAY-1999; 99US-0134256.
XX
XX 14-MAY-1999; 99US-0134218.
XX
XX 14-MAY-1999; 99US-0134219.
XX
XX 14-MAY-1999; 99US-0134221.
XX
XX 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143544.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.

```

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

```

```

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 63.4%; Score 45; DB 21; Length 682;
Best Local Similarity 69.2%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 3 LASFLPLNSFL 15
Db 70 LASSFLWIGNSFL 82

```

```

RESULT 3
ABB92235
ID ABB92235 standard; Protein; 682 AA.

```

```

AC ABB92235;
DT 31-MAY-2002 (first entry)

```

```

DE Herbicidally active polypeptide SEQ ID NO 1446.

```

```

XX Herbicidal; plant; agriculture; herbicide.

```

```

XX Arabidopsis thaliana.

```

```

XX WO200210210-A2.

```

```

XX 07-FEB-2002.

```

```

XX 28-AUG-2001; 2001WO-EP09892.

```

```

XX 28-AUG-2001; 2001WO-EP09892.

```

```

XX (FARB ) BAYER AG.

```

```

XX Tierjen K, Weidler M;

```

```

XX WPI; 2002-269010/31.

```

```

PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -

```

```

XX Claim 5; SEQ ID NO 1446; 261pp + Sequence listing; English.

```

```

XX The invention relates to identifying target proteins
XX (ABB90790-ABB94016) for herbicidally active compounds, comprising

```

```

CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.

```

```

XX Sequence 682 AA;

```

```

Query Match 63.4%; Score 45; DB 23; Length 682;
Best Local Similarity 69.2%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 3 LASFLPLNSFL 15
Db 70 LASSFLWIGNSFL 82

```

RESULT 4
AAG38851
ID AAG38851 standard; Protein; 725 AA.
XX
AC AAG38851;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47990.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

```

PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159367.
PR 14-OCT-1999; 99US-0159368.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match      63.4%; Score 45; DB 21; Length 725;
Best Local Similarity 69.2%; Pred. No. 13;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 3 LASEPLFLNSFL 15
DB 70 LASSPLWIGNSFL 82

RESUR 5
ABP03996
ID ABP03996 standard; Protein; 60 AA.
XX
XX ABP03996;
XX
XX 24-JUN-2002 (first entry)

```

```

XX DE Human ORFX protein sequence SEQ ID NO:7974.
XX KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
XX KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
XX KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX KW hypertension; hypothyroidism; cholesterol ester storage disease;
XX KW immune deficiency; immune disorder; infectious disease;
XX KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
XX KW myasthenia gravis.
XX OS Homo sapiens.
XX MO200192523-A2.
XX PD 06-DEC-2001.
XX PF 29-MAY-2001; 2001MO-US10836.
XX PR 30-MAY-2000; 2000US-206132P.
XX PR 29-AUG-2000; 2000US-228716P.
XX PA (CURA-) CURAGEN CORP.
XX PI Shinkens RA, Leach MD;
XX DR WPI; 2002-106308/14.
XX DR N-PSDB; ABN19748.
XX PS
XX PT Novel human polypeptides and polynucleotides useful for diagnosing,
XX PT preventing and treating cardiovascular disease, neurodegenerative,
XX PT hyperproliferative disorders and autoimmune disorders
XX PS
XX PS Disclosure; SEQ ID 7974; 1037bp; English.
XX
XX CC The present invention describes substantially purified human proteins
XX CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
XX CC in the specification). ABN15762 to ABN27252 encode the human ORFX
XX CC proteins given in ABP0010 to ABP11500. ORFX proteins are useful for
XX CC treating or preventing a pathology associated with an ORFX-associated
XX CC disorder in humans, and in the manufacture of a medicament for treating a
XX CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
XX CC sequences can be used in gene therapy. ORFX sequences can be used in the
XX CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
XX CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
XX CC osteoarthritis, neurodegenerative disorders, disorders related to organ
XX CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
XX CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
XX CC storage disease, various immune deficiencies and disorders, infectious
XX CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
XX CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
XX CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
XX CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
XX CC bone degenerative disorders, or periodontal disease, and for gut
XX CC protection or regeneration and treatment of lung or liver fibrosis,
XX CC reperfusion injury in various tissues and conditions resulting from
XX CC systemic cytokine damage.
XX CC N.B. The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 60 AA;

```

```

Query Match      59.2%; Score 42; DB 23; Length 60;
Best Local Similarity 50.0%; Pred. No. 3.2;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
OY 1 IDLASEPLFLNSFL 14
DB 37 VILPCSFIFINSF 50

```

```

RESULT 6
AAV55732
ID AAV55732 strand; Protein; 106 AA.
XX
AC AAV55732;
XX
DT 22-FEB-2000 (first entry)
XX
DE FKHL5 forkhead domain fragment.
XX
KW Forkhead transcription factor gene; FKHL7; treatment; gliucoma; human;
XX transgenic animal; drug screening.
XX
OS Unidentified.
XX
PN WO9953060-A2.
XX
PD 21-OCT-1999.
XX
PF 14-APR-1999; 99WO-US08148.
XX
PR 15-APR-1998; 98US-0081870.
XX PR 22-MAY-1998; 98US-0083352.
XX
PA (IOWA ) UNIV IOWA RES FOUND.
XX
PI Sheffield VC, Alward WLM, Stone EM, Nishimura D, Patil S;
XX WPI; 1999-620429/53.
XX
DR WPI; 1999-620429/53.
XX
PT New isolated human forkhead transcription factor gene, FKHL7, used to,
XX e.g. developt products for the diagnosis -
XX
PS Disclosure; Fig 2; 99pp; English.
XX
CC The invention provides a human forkhead transcription factor gene,
XX FKHL7. The FKHL7 protein can be produced by standard recombinant
XX methodology. The products can be used for diagnosis, prognosis,
XX monitoring, prevention and treatment of glaucoma. They can also be used
XX for the production of transgenic animals and drug screening. Sequences
XX CC AAV5722-739 represent forkhead domains of different members of the
XX CC FKHL-family of genes.
XX
SQ Sequence 106 AA;
XX
Query Match 59.2%; Score 42; DB 20; Length 106;
Best Local Similarity 64.3%; Pred. No. 5.9;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IDLASEFLFLSNF 14
Db 82 IDPASEFMENGSP 95

RESULT 7
AAV55733
ID AAV55733 strand; Protein; 106 AA.
XX
AC AAV55733;
XX
DT 22-FEB-2000 (first entry)
XX
DE FKHL6 forkhead domain fragment.
XX
KW Forkhead transcription factor gene; FKHL7; treatment; glaucoma; human;
XX transgenic animal; drug screening.
XX
OS Unidentified.
XX
PN WO9953060-A2.
XX
PD 21-OCT-1999.
XX

```

```

PF 14-APR-1999; 99WO-US08148.
XX
PR 15-APR-1998; 98US-0081870.
XX PR 22-MAY-1998; 98US-0083352.
XX
PA (IOWA ) UNIV IOWA RES FOUND.
XX
PI Sheffield VC, Alward WLM, Stone EM, Nishimura D, Patil S;
XX WPI; 1999-620429/53.
XX
PT New isolated human forkhead transcription factor gene, FKHL7, used to,
XX e.g. developt products for the diagnosis -
XX
PS Disclosure; Fig 2; 99pp; English.
XX
CC The invention provides a human forkhead transcription factor gene,
XX FKHL7. The FKHL7 protein can be produced by standard recombinant
XX methodology. The products can be used for diagnosis, prognosis,
XX monitoring, prevention and treatment of glaucoma. They can also be used
XX for the production of transgenic animals and drug screening. Sequences
XX CC AAV5722-739 represent forkhead domains of different members of the
XX CC FKHL-family of genes.
XX
SQ Sequence 106 AA;
XX
Query Match 59.2%; Score 42; DB 20; Length 106;
Best Local Similarity 64.3%; Pred. No. 5.9;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IDLASEFLFLSNF 14
Db 82 IDPASEFMENGSP 95

RESULT 8
ABG03158
ID ABG03158 strand; Protein; 465 AA.
XX
AC ABG03158;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #3149.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX DR N-PSDB; AAS67345.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
PS Claim 20; SEQ ID NO 33517; 103pp; English.
XX

```


CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridization probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
XX
SQ Sequence 465 AA;
XX
Query Match 57.7%; Score 41; DB 22; Length 465;
Best Local Similarity 64.3%; Pred. No. 43;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 IDLASEFLPLNSNF 14
DB 124 IDPASEFMEEGSF 137
XX
RESULT 9
ID AAU19348
AAU19348 standard; Protein; 183 AA.
XX
AC AAU19348;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human G protein-coupled receptor nGPR-2274.
XX
XX Human; G protein-coupled receptor; nGPR-x; antiviral; analgesic;
KW cytosolic; cardiac; antidiabetic; anorectic; hypotensive; hyperensive;
KW antiapoptotic; neurotrophic; neuroprotective; antidepressant;
KW viral infection; HIV-1; human immunodeficiency virus; HIV-2; pain;
KW cancer; metabolic disease; cardiovascular disease; type 2 diabetes;
KW obesity; anorexia; hypotension; hypertension; myocardial infarction;
KW atherosclerosis; Parkinson's disease; psychosis; neurological disorder;
KW schizophrenia; migraine; major depression; anxiety; mental disorder;
KW manic depression; dyskinesia; Huntington's disease; Tourette's Syndrome.
XX
OS Homo sapiens.
XX
XX MO200166751-A2.
XX
XX 13-SEP-2001.
XX
XX PD 08-MAR-2001; 2001WO-US07370.
XX
XX PF 08-MAR-2000; 2000US-0187583.
XX
XX PR 08-MAR-2000; 2000US-0187584.
XX
XX PR 08-MAR-2000; 2000US-0187637.
XX
XX PR 08-MAR-2000; 2000US-0187639.
XX
XX PR 08-MAR-2000; 2000US-0187640.
XX
XX PR 08-MAR-2000; 2000US-0187707.
XX
XX PR 08-MAR-2000; 2000US-0187708.
XX
XX PR 08-MAR-2000; 2000US-0187709.
XX
XX PR 08-MAR-2000; 2000US-0187827.
XX
XX PR 08-MAR-2000; 2000US-0188290.
XX
XX PA (PHMA) PHARMACIA & UPJOHN CO.

XX
PI Vogel; G. Deluca M;
XX
DR MPI; 2001-536779/59.
DR N-PSDB; AAS30920.
XX
PT Isolated nucleic acid molecules encoding G protein-coupled receptors
PT termed nGPR-x, useful in the treatment and diagnosis of viral
PT infections, cancers and mental disorders (e.g. Parkinson's disease and
PT schizophrenia) -
XX
PS Claim 31; Page 250-251; 292pp; English.
XX
CC The invention relates to novel isolated nucleic acid molecules encoding
CC G protein-coupled receptors, termed nGPR-x, nGPR-x polynucleotides,
CC polypeptides, and modulators may be used in the treatment of diseases and
CC conditions such as infections, such as viral infections caused by HIV-1
CC (human immunodeficiency virus) or HIV-2, pain, cancers, metabolic and
CC cardiovascular diseases and disorders (e.g., type 2 diabetes, obesity,
CC anorexia, hypotension, hypertension, myocardial infarction,
CC atherosclerosis), Parkinson's disease, and psychotic and
CC neurological disorders, including schizophrenia, migraine, major
CC depression, anxiety, mental disorder, manic depression, and
CC dyskinesias, such as Huntington's disease or Tourette's Syndrome
CC and many other diseases and syndromes listed in the specification.
CC nGPR-x polynucleotides and polypeptides, as well as nGPR-x
CC modulators, may also be used in diagnostic assays for such diseases or
CC conditions. The present sequence represents a G protein-coupled receptor
CC of the invention.
XX
SQ Sequence 183 AA;
XX
Query Match 56.3%; Score 40; DB 22; Length 183;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 FLPLNSNF 14
DB 58 FLPLNSNF 65
XX
RESULT 10
ID ABP30033
ABP30033 standard; Protein; 280 AA.
XX
AC ABP30033;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 9242.
XX
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
XX Streptococcus pyogenes.
XX
XX OS
XX
XX MO200234771-A2.
XX
XX PN 02-MAY-2002.
XX
XX PD 29-OCT-2001; 2001WO-GB04789.
XX
XX PF 27-OCT-2000; 2000GB-0026333.
XX
XX PR 24-NOV-2000; 2000GB-0028727.
XX
XX PR 07-MAR-2001; 2001GB-0005640.
XX
XX PA (CHIR-) CHIRON SPA.
XX
XX PA (GENO-) INST GENOMIC RES.
XX
XX PI Telford J, Maignant V, Margaric Ros YI, Grandi G, Fraser C;
XX PI Tettein H;

DR	WPI, 2002-352536/38.
DR	N-PSDB; AEN70664.
XX	
PT	New Streptococcus protein for the treatment or prevention of infection
PT	or disease caused by Streptococcus bacteria, such as meningitis, and
PT	for detecting a compound that binds to the protein -
XX	
PS	Claim 1; Page 4048; 4525bp; English.
XX	
CC	The invention relates to a protein (ABP25413-ABP30895) from group B
CC	streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC	(Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC	the specification. The proteins have antibacterial and antiinflammatory
CC	activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and
CC	antibodies that bind (1) are used in the manufacture of medicaments for
CC	the treatment or prevention of infection or disease caused by
CC	Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC	Nucleic acids encoding (1) are used to detect Streptococcus in a
CC	biological sample. (1) is used to determine whether a compound binds to
CC	(1). A composition comprising (1) or a nucleic acid encoding (1), may be
CC	used as a vaccine or diagnostic composition. The disease caused by
CC	Streptococcus that is prevented or treated may be meningitis. Nucleic
CC	acid encoding (1) may be used to recombinantly produce (1) and may be
CC	used in gene therapy. Antibodies to (1) are used for affinity
CC	chromatography, immunoassays, and distinguishing/identifying
CC	Streptococcus proteins.
CC	
SQ	Sequence 280 AA;
XX	
Query Match	56.3%; Score 40; DB 23; Length 280;
Best Local Similarity	77.8%; Pred. No. 38;
Matches 7; Conservative	2; Mismatches 0; Indels 0; Gaps 0;
OY	6 EFLFSLNSF 14 :- :-
Db	59 EFLFSLNSF 67
RESULT 11	
ID	AAU14133
AC	AAU14133 standard; Protein; 554 AA.
XX	
DT	24-OCT-2001 (first entry)
XX	
DE	Human novel protein #4.
XX	
KW	Human; novel protein; Antianaemic; osteopathic; antiinflammatory;
KW	immunomodulatory; cyostatic; neuroprotective; vulnerrary; nootropic;
KW	anticonvulsant; antiallergic; cerebroprotective; antifungal; antiviral;
KW	antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
KW	chromolytic; immunogen; antibody; gene therapy; neurological disorder;
KW	Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW	tissue regeneration; immune disorder.
XX	
OS	Homo sapiens.
XX	
PN	WO200155437-A2.
XX	
PD	02-AUG-2001.
XX	
PF	25-JAN-2001; 2001WO-US02623.
XX	
PR	25-JAN-2000; 2000US-0491404.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Drmanac RT;
XX	
DR	WPI; 2001-451939/48.
DR	N-PSDB; AAS22438.
XX	

PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216547.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225477.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228824.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 29-SEP-2000; 2000US-0235636.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0244675.
PR 08-NOV-2000; 2000US-0244676.
PR 08-NOV-2000; 2000US-0244677.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251859.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259676.

(HUMA-) HUMAN GENOME SCI INC.
XX PA
XX PI
XX Rosen CA, Barash SC, Ruben SW;
XX MPI: 2001-483426/52.
XX DR N-PSDB; AAK63634.
XX DR
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX PT useful for preventing, diagnosing and/or treating cancers and
XX PT metastasis -
XX XX
XX PS Clatm 11; SEQ ID NO 18446; 3071bp + Sequence Listing; English.
XX CC AAK4951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
XX CC activity, and can be used in gene therapy and vaccine production. (I)
XX CC proteins and polynucleotides may be used in the prevention, diagnosis and
XX CC treatment of diseases associated with inappropriate (I) expression. For
XX CC example, they may be used to treat disorders associated with decreased

expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/haematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAM82169 represent sequences used in the exemplification of the present invention.

Sequence 53 AA;

Query Match 53.5%; Score 38; DB 22; Length 53;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
8 LFLSNSFL 15
7 LFLSNSFL 14
Db

RESULT 13
ABU00902
ID ABU00902 standard; Protein; 134 AA.

AC ABU00902;

DT 11-FEB-2003 (first entry)

DE S. pneumoniae type 4 strain protein from coding region #471.

KM Bacterial meningitis; pneumonia; sepsis; otitis media;
KM ear infection; antiinflammatory; antibacterial; immunostimulant;
KM auditory; respiratory; gene therapy; vaccine.

OS Streptococcus pneumoniae type 4 strain.

PN WO200277021-A2.

PD 03-OCT-2002.

PF 27-MAR-2002; 2002WO-IB02163.

PR 27-MAR-2001; 2001GB-0007658.

PA (CHIR-) CHIRON SPA.

PA (GENO-) INSR GENOMIC RES.

PI Masignani V, Tettelin H, Fraser C;

DR WPI, 2003-040579/03.

DR N-PSDB; ABX06183.

PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
PT useful as medicaments for treating or preventing a disease or infection
PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media
PT or ear infection

PI Claim 1; SEQ ID No 942; 56pp; English.

The invention relates to a protein comprising or having at least 50% identity to any of the 2469 amino acid sequences, identified in the specification (available on a computer readable format), or its fragment, expressed from 2469 of 2489 identified DNA coding regions from the Streptococcus pneumoniae type 4 strain genomic sequence appearing as AAB56454. Also included are an antibody which binds one of the proteins, treating a patient by administering the protein, DNA or antibody (in a composition), a kit comprising first and second primers, which are the nucleic acid cited above or fragments between nucleotides 8-100 of a sequence not defined in the specification, for amplifying a target sequence contained within a Streptococcus nucleic acid sequence,

where the first primer is substantially complementary to the target CC sequence and the second primer is substantially complementary to the CC complement of the target sequence, and where the parts of the primers CC having substantial complementarity define the termini of the target CC sequence to be amplified, assay comprising contacting a test compound CC with the protein, and determining whether the test compound binds to the CC protein and a Streptococcus pneumoniae bacterium, where one or more CC genes encoding the proteins has been rendered inactive. The proteins, CC nucleic acid molecules, antibody and compositions are useful as CC medicaments for treating or preventing a disease or infection due to CC streptococcus bacteria, particularly S. pneumoniae, such as pneumonia, CC sepsis, otitis media or ear infection. They are also useful in developing CC vaccines, diagnostics and antibiotics. The methods are useful for CC identifying immunodominant proteins. The present sequence is one of CC the 2469 proteins expressed by the identified coding regions from the CC genomic sequence.

Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in electronic
format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences.

Sequence 134 AA;

Query Match 53.5%; Score 38; DB 24; Length 134;
Best Local Similarity 61.5%; Pred. No. 40;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

3 LASEPLFLSNSFL 15

24 LVSTYFLSRQFL 36

Db

RESULT 14
ABG05253
ID ABG05253 standard; Protein; 253 AA.

AC ABG05253;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #5244.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAB69440.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PI Claim 20; SEQ ID No 35612; 103pp; English.

The invention relates to isolated polynucleotide (I) and
polypeptide (II) sequences. (I) is useful as hybridisation probes,
polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

Query Match 53.5%; Score 38; DB 22; Length 253;
 Best Local Similarity 50.0%; Pred. No. 78;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 DIASEFLFLNSFL 15
 DB 20 DFLTEYLFVMSFL 33

RESULT 15

ABB63044
 ID ABB63044 standard; Protein; 467 AA.

AC ABB63044;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 15924.

KM Drosophila developmental biology; cell signalling; insecticide;
 KM pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

DR MPI; 2001-656860/75.

DR N-PGDB; ABL07147.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS Disclosure; SEQ ID NO 15924; 21bp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

Search completed: November 26, 2003, 12:30:26
 Job time : 42.5663 secs

Query Match 53.5%; Score 38; DB 22; Length 467;
 Best Local Similarity 57.1%; Pred. No. 1.5e+02;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 DIASEFLFLNSFL 15
 DB 251 DVNSEMLWYQNSFL 264

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 26, 2003, 12:22:05 ; Search time 7.22892 Seconds

(without alignments)
97.580 Million cell updates/sec

Title: US-09-230-111c-15
Sequence: 1 IDLASEFLFLSNSFL 15

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	42	59.2	98	1 YAM3_YEAST
2	41	57.7	353	1 FXPI_MOUSE
3	41	57.7	354	1 FXPI_HUMAN
4	41	57.7	444	1 FXPI_HUMAN
5	41	57.7	1239	1 DPGI_MOUSE
6	40	56.3	1401	1 WIS4_SCHPO
7	39	54.9	72	1 NULM_COTJA
8	39	54.9	99	1 NULM_MERTSE
9	39	54.9	165	1 LSPA_MERTSE
10	39	54.9	409	1 NULM_CABEL
11	38	53.5	283	1 SURE_CHLMU
12	38	53.5	456	1 TRME_LEPIN
13	38	53.5	735	1 YQJ7_SCHPO
14	38	53.5	1816	1 AFE_HUMAN
15	37	52.1	245	1 YCBF_BACSU
16	37	52.1	324	1 NULM_CHICK
17	37	52.1	411	1 CARA_YEAST
18	37	52.1	503	1 LEUI_BUCUM
19	37	52.1	1096	1 PDPI_ARATH
20	37	52.1	1178	1 MNM4_YEAST
21	36	50.7	1	1 ILIB_CANFA
22	36	50.7	164	1 LSPA_MOUSE
23	36	50.7	164	1 LSPA_MOUSE
24	36	50.7	164	1 LSPA_MOUSE
25	36	50.7	172	1 YHBN_HAEIN
26	36	50.7	293	1 YQFU_BACSU
27	36	50.7	346	1 YGXA_BACSU
28	36	50.7	346	1 GATD_ECOLI
29	36	50.7	383	1 FLHB_BUCAL
30	36	50.7	502	1 C932_SOYBN
31	36	50.7	510	1 C933_SOYBN
32	36	50.7	517	1 EUGI_YEAST
33	36	50.7	1375	1 BNRI_YEAST

34	35.5	50.0	516	1 YJUI_ECOLI	P37342 escherichia
35	35	49.3	251	1 YPIG_BACSU	P35154 bacillus su
36	35	49.3	308	1 GLS2_ECOLI	P77470 escherichia
37	35	49.3	380	1 METC_LACLA	O91859 lactococcus
38	35	49.3	399	1 PDK_ASCSU	O02623 ascaris suu
39	35	49.3	451	1 ARGD_ALINGL	O04866 alnus gluti
40	35	49.3	493	1 ARGD_ALINGL	O9ca71 arabidopsis
41	35	49.3	502	1 LEUI_BUCUM	O9ev74 buchneia ap
42	35	49.3	510	1 CP5V_CANNA	O12587 candida mai
43	35	49.3	519	1 CP5V_CANNA	P43083 candida api
44	35	49.3	519	1 CP5V_CANNA	O12573 candida api
45	35	49.3	670	1 NUSM_WHEAT	O37680 triticum ae

ALIGNMENTS

RESULT 1	YAM3_YEAST	STANDARD;	PRT;	98 AA.
AC	P39559;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	01-NOV-1995 (Rel. 32, Last annotation update)			
DE	Hypothetical 11.1 kDa protein in FL01-PH01 intergenic region.			
GN	VAR053W.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.			
OX	NCBI_TaxID=4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=S288c / AB972;			
RC	MEDLINE=95249563; PubMed=7731988;			
RX	Busey H., Kaback D.B., Zhong W., Vo D.T., Clark M.W., Fortin N.,			
RA	Hall J., Ouellette B.F.F., Keng T., Barton A.B., Su Y., Davies C.K.,			
RT	Storms R.K.;			
RT	"The nucleotide sequence of chromosome I from Saccharomyces			
RT	cerevisiae."			
RL	Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813 (1995).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL; L28920; AAC09500.1; -			
DR	PIR; S53467; S53467.			
DR	SGD; S0000085; YAR053W.			
KW	Hypothetical protein; Transmembrane.			
FT	TRANSMEM 13			
FT	TRANSMEM 33			
FT	TRANSMEM 65			
FT	TRANSMEM 85			
FT	TRANSMEM 98			
FT	SEQUENCE 98 AA; 11143 MW; 4A8D7D8B17BC9832 CRC64;			
QY	Query Match			
QY	Best local similarity 59.2%; Score 42; DB 1; Length 98;			
QY	Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;			
DB	1 IDLASEFLFLSNSFL 14			
DB	75 VNLPCSFIFISNSF 88			
RESULT 2				

```

FXFL MOUSE
ID FXFL_MOUSE STANDARD; PRT; 353 AA.
AC 061080; 061661;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Forkhead box protein F1 (Forkhead-related protein FKHL5) (Forkhead-
DE related transcription factor 1) (FREC-1) (Hepatocyte nuclear factor 3
DE forkhead homolog 8) (HFH-8).
GN FOXF1 OR FOXF1A OR FKHL5 OR FREAC1 OR HFH8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=96224034; PubMed=8626802;
RA Hellqvist M., Mahlapuu M., Samuelsson L., Enerbaeck S., Carlsson P.;
RT "Differential activation of lung-specific genes by two forkhead
RT proteins, FREAC-1 and FREAC-2."
RL J. Biol. Chem. 271:4482-4490(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=95046902; PubMed=7958446;
RA Clevidence D.E., Overdier D.G., Peterson R.S., Porcella A., Ye H.,
RA Paulson K.E., Costa R.H.;
RT "Members of the HNF-3/forkhead family of transcription factors
RT exhibit distinct cellular expression patterns in lung and regulate
RT the surfactant protein B promoter."
RL Dev. Biol. 166:195-209(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=21213904; PubMed=11313147;
RA Chang V.W.H., Ho Y.S.;
RT "Structural characterization of the mouse Foxfla gene."
RL Gene 267:201-211(2001).
CC -1- FUNCTION: PROBABLE TRANSCRIPTION ACTIVATOR FOR A NUMBER OF LUNG-
CC SPECIFIC GENES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: LUNG AND PLACENTA.
CC -1- DOMAIN: ACTIVATION DOMAINS C-TERMINAL OF (AND DISTINCT FROM) THE
CC FORKHEAD DOMAINS ARE NECESSARY FOR TRANSCRIPTIONAL ACTIVATION (BY
CC SIMILARITY).
CC -1- SIMILARITY: Contains 1 fork-head domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U42556; AAC2445.1; -.
CC EMBL: U35949; AAG4885.1; AUT_INIT.
CC EMBL: AF346834; AAK35051.1; -.
CC HSSP: O63245; 2HFH.
CC TRANSFAC: T02461; -.
CC MGD: MGI:1347470; Foxfla.
CC GO: GO:0030324; P:lung development; IMP.
CC GO: GO:0009887; P:organogenesis; IMP.
CC GO: GO:0001570; P:vasculogenesis; IMP.
CC InterPro: IPR001766; TF_Fork_head.
CC Pfam: PR00250; Fork_head; 1.
CC PRINTS: PR00053; FORKHEAD.
CC ProDom: PD000425; TF_Fork_head; 1.
CC SMART: SM00339; FH; 1.
CC PROSITE: PS00657; FORK_HEAD_1; 1.
CC PROSITE: PS00658; FORK_HEAD_2; 1.
CC PROSITE: PS50039; FORK_HEAD_3; 1.

```

```

KM DNA-Binding; Nuclear protein; Transcription regulation; Activator.
FT DNA BIND 22 113
FT CONFLICT 7 19
FT CONFLICT 207 247
FT CONFLICT 207 247
FT CONFLICT 207 247
FT CONFLICT 207 247
SQ SEQUENCE 353 AA; 37798 MW; 046AE08D1765A69 CRC64;
Query Match 57.7%; Score 41; DB 1; Length 353;
Best Local Similarity 64.3%; Pred. No. 5.3;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 IDLASEFLPLNSNF 14
Db 99 IDPASEFMEEGSF 112
RESULT 3
FXFL_HUMAN
ID FXFL_HUMAN STANDARD; PRT; 354 AA.
AC Q12946;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Forkhead box protein F1 (Forkhead-related protein FKHL5) (Forkhead-
DE related transcription factor 1) (FREC-1) (Forkhead-related activator-
DE 1).
GN FOXF1 OR FKHL5 OR FREAC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=96224034; PubMed=8626802;
RA Hellqvist M., Mahlapuu M., Samuelsson L., Enerbaeck S., Carlsson P.;
RT "Differential activation of lung-specific genes by two forkhead
RT proteins, FREAC-1 and FREAC-2."
RL J. Biol. Chem. 271:4482-4490(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98443197; PubMed=9769171;
RA Mahlapuu M., Pelto-Huikko M., Aitola M., Enerbaeck S., Carlsson P.;
RT "FREAC-1 contains a cell-type-specific transcriptional activation
RT domain and is expressed in epithelial-mesenchymal interfaces."
RL Dev. Biol. 202:183-195(1998).
RN [3]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=95045392; PubMed=7957066;
RA Pierron S., Hellqvist M., Samuelsson L., Enerbaeck S., Carlsson P.;
RT "Cloning and characterization of seven human forkhead proteins:
RT binding site specificity and DNA bending."
RL EMBO J. 13:5002-5012(1994).
CC -1- FUNCTION: PROBABLE TRANSCRIPTION ACTIVATOR FOR A NUMBER OF LUNG-
CC SPECIFIC GENES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: LUNG AND PLACENTA.
CC -1- DOMAIN: ACTIVATION DOMAINS C-TERMINAL OF (AND DISTINCT FROM) THE
CC FORKHEAD DOMAINS ARE NECESSARY FOR TRANSCRIPTIONAL ACTIVATION.
CC -1- SIMILARITY: Contains 1 fork-head domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U13219; AAC50399.1; -.
CC EMBL: AF085343; AAC61576.1; -.

```



```

DR EMBL: AF085342; AAC61576.1; JOINED.
DR HSSP: 063245; 2HPH.
DR TRANSFAC: T02464; -.
DR Genew; HGNC:3809; FOXF1.
DR MIM; 601089; -.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0005667; C:transcription factor complex; TAS.
DR GO; GO:0005515; P:protein binding activity; TAS.
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR SMART; SM00325; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS0039; FORK_HEAD_3; 1.
DR DNA-binding; Nuclear protein; Transcription regulation; Activator.
KW DNA_BIND
SQ SEQUENCE 354 AA; 38034 MW; DC32373A7AD034EF CRC64;

Query Match 57.7%; Score 41; DB 1; Length 354;
Best Local Similarity 64.3%; Pred. No. 5.3;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 1 IDLASEPFLNSPF 14
Db 99 IDPASEFWFEEGSF 112

RESULT 4
FXF2_HUMAN STANDARD; PRT; 444 AA.
ID Q12947; Q9U085;
AC Q12947; Q9U085;
DT 01-NOV-1997 (Rel. 35; Created)
DT 30-MAY-2000 (Rel. 39; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Forkhead box protein F2 (Forkhead-related protein FKHL6) (Forkhead-related transcription factor 2) (FREC-2) (Forkhead-related activator-2).
GN FOXF2 OR FKHL6 OR FREAC2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OC NCBI_TaxId:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99017977; PubMed=9799607;
RA Blixt A., Mahlapuu M., Bjurell C., Darntofe C., Johannesson T., Enerbaeck S., Carlsson P.;
RT "The two-exon gene of the human forkhead transcription factor FREAC-2 (FKHL6) is located at 6p25.3."
RT Genomics 53:387-390(1998).
RN [2]
RP SEQUENCE OF 37-444 FROM N.A.
RC TISSUE=lung;
RX MEDLINE=96224034; PubMed=8626802;
RA Hellqvist M., Mahlapuu M., Samuelsson L., Enerbaeck S., Carlsson P.;
RT "Differential activation of lung-specific genes by two forkhead proteins, FREAC-1 and FREAC-2."
RT J. Biol. Chem. 271:4482-4490(1996).
RN [3]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=95045392; PubMed=7957066;
RA Pieterou S., Hellqvist M., Samuelsson L., Enerbaeck S., Carlsson P.;
RT "Cloning and characterization of seven human forkhead proteins: binding site specificity and DNA bending."
RT EMBO J. 13:5002-5012(1994).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=9889768; PubMed=9722567;
RA Hellqvist M., Mahlapuu M., Blixt A., Enerbaeck S., Carlsson P.;
RT "The human forkhead protein FREAC-2 contains two functionally redundant activation domains and interacts with TBP and TFIIB."

```

```

RL J. Biol. Chem. 273:23335-23343(1998).
CC -1- FUNCTION: PROBABLE TRANSCRIPTION ACTIVATOR FOR A NUMBER OF LUNG-SPECIFIC GENES.
CC -1- SUBUNIT: INTERACTS WITH THE TRANSCRIPTION FACTORS TBP AND TFIIB.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: LUNG AND PLACENTA.
CC -1- DOMAIN: TWO ACTIVATION DOMAINS, AD1 AND AD2, C-TERMINAL OF (AND DISTINCT FROM) THE FORKHEAD DOMAINS ARE NECESSARY FOR TRANSCRIPTIONAL ACTIVATION.
CC -1- SIMILARITY: Contains 1 fork-head domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF084939; AAD19875.1; JOINED.
DR EMBL: AF084938; AAD19875.1; JOINED.
DR EMBL: U13220; AAC32226.1; -.
DR PIR; T09474; T09474.
DR HSSP; 063245; 2HPH.
DR TRANSFAC: T02465; -.
DR Genew; HGNC:3810; FOXF2.
DR MIM; 603250; -.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0005667; C:transcription factor complex; TAS.
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR Prodom; PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS0039; FORK_HEAD_3; 1.
KW DNA-binding; Nuclear protein; Transcription regulation; Activator.
FT DOMAIN 33 41 POLY-ALA.
FT DOMAIN 46 51 POLY-SER.
FT DOMAIN 58 67 POLY-SER.
FT DOMAIN 77 84 POLY-GLY.
FT DNA_BIND 99 190 FORK-HEAD.
FT DOMAIN 262 272 POLY-HIS.
FT DOMAIN 301 306 POLY-GLY.
FT DOMAIN 312 315 POLY-SER.
FT DOMAIN 429 432 POLY-HIS.
SQ SEQUENCE 444 AA; 45993 MW; 32BDC5F373CFB147 CRC64;

Query Match 57.7%; Score 41; DB 1; Length 444;
Best Local Similarity 64.3%; Pred. No. 6.7;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 1 IDLASEPFLNSPF 14
Db 176 IDPASEFWFEEGSF 189

RESULT 5
IDL_MOUSE STANDARD; PRT; 1239 AA.
ID P54059; Q9U128;
AC P54059; Q9U128;
DT 01-OCT-1996 (Rel. 34; Created)
DT 28-FEB-2003 (Rel. 41; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE DNA polymerase gamma subunit 1 (EC 2.7.7.7) (Mitochondrial DNA polymerase catalytic subunit) (POLG-alpha).
GN POLG OR POLG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxId:10090;
RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN-BALB/c; TISSUE=Muscle;
RA Chang S.W., Colvin S., Sarkos P., Denniger G., Zassenhaus H.P.;
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=20514044; PubMed=11058962;
RA Mort J.L., Denniger G., Zullo S.J., Zassenhaus H.P.;
RT "Genomic structure of murine mitochondrial DNA polymerase-gamma.";
RL DNA Cell Biol. 19:601-605 (2000).
CC -1- FUNCTION: INVOLVED IN THE REPLICATION OF MITOCHONDRIAL DNA.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA] (N).
CC -1- COFACTOR: MAGNESIUM (By similarity).
CC -1- SUBUNIT: Heterotrimer composed of a catalytic subunit and an
CC homodimer of accessory subunits.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, U53584; AAF98977.1; -
DR EMBL, AF268975; AAF82772.1; -
DR EMBL, AF268970; AAF82772.1; JOINED.
DR EMBL, AF268971; AAF82772.1; JOINED.
DR EMBL, AF268972; AAF82772.1; JOINED.
DR EMBL, AF268973; AAF82772.1; JOINED.
DR EMBL, AF268974; AAF82772.1; JOINED.
DR MGD; MGI:1196389; Polg.
DR InterPro; IPR001098; DNA pol.
DR InterPro; IPR002297; DNA polG.
DR Pfam; PF00476; DNA pol A.1.
DR PRINTS; PR00867; DNAPOLG.
DR SMART; SM00482; POLAC.1.
DR PROSITE; PS00447; DNA_POLYMERASE_A.1.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Mitochondrion; Magnesium.
FT DOMAIN 517 520
FT POLY-GLU
FT CONFLICT 23 24 AR -> EG (IN REF. 1).
FT CONFLICT 27 27 A -> S (IN REF. 1).
FT CONFLICT 32 32 D -> A (IN REF. 1).
FT CONFLICT 159 159 T -> N (IN REF. 1).
FT CONFLICT 228 228 I -> T (IN REF. 1).
FT CONFLICT 303 308 HKNPAA -> QODPAV (IN REF. 1).
FT CONFLICT 324 324 E -> D (IN REF. 1).
FT CONFLICT 327 327 MISSING (IN REF. 1).
FT CONFLICT 441 454 EMKSLMDLANDAC -> DDERVVVWMLMPG (IN REF.
FT 1).
FT CONFLICT 557 557 O -> K (IN REF. 1).
FT CONFLICT 666 666 N -> S (IN REF. 1).
FT CONFLICT 861 861 P -> T (IN REF. 1).
FT CONFLICT 886 886 F -> L (IN REF. 1).
FT CONFLICT 919 921 GIS -> RH (IN REF. 1).
FT CONFLICT 1024 1024 A -> S (IN REF. 1).
FT CONFLICT 1233 1233 H -> Y (IN REF. 1).
SQ SEQUENCE 1239 AA; 138898 MW; 77E514FA0FB679PE CRC64;
Query Match 57.7%; Score 41; DB 1; Length 1239;
Best Local Similarity 61.5%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

RESULT 6
ID WIS4 SCHPO
AC WIS4 SCHPO STANDARD; PRT; 1401 AA.
DT 014259; P87062; Q92384;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE MAP kinase kinase kinase wisk5 (EC 2.7.1.-) (MAP kinase kinase kinase
DE wisk) (MAP kinase kinase kinase wisk).
GN WIS4 OR WAK1 OR WIK1 OR SPK9G1.02.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98026892; PubMed=9321395;
RA Samejima I., Mackie S., Fantes P.A.;
RT "Multiple modes of activation of the stress-responsive MAP kinase
RT pathway in fission yeast.";
RL EMO J. 16:6162-6170 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Haylee J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor J., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Gymnopoulos B.,
RA Weljens J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel K., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu S., Dreano S., Gloux S., Leleaux V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurt S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garcon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880 (2002).
RN [3]
RP SEQUENCE OF 96-1401 FROM N.A.
RC STRAIN=972;
RX MEDLINE=97282620; PubMed=9136929;
RA Shieh J.C., Wilkinson M., Buck V., Morgan B., Makino K.,
RA Miller J.B.A.;
RT "The Mca4 response regulator coordinately controls the stress-
RT activated wisk1-Wisk1-Styl1 MAP kinase pathway and fission yeast cell
RT cycle.";
RL Genes Dev. 11:1008-1022 (1997).
RN [4]
RP SEQUENCE OF 457-543 FROM N.A.
RC STRAIN=972;
RX MEDLINE=97331762; PubMed=918094;
RA Shiozaki K., Shiozaki M., Russell P.;
RT "Mca4 mitotic catastrophe suppressor regulates the fission yeast cell
RT cycle through the wisk1-Wisk1-Styl1 kinase cascade.";
RL Mol. Biol. Cell 8:409-419 (1997).
CC -1- FUNCTION: INVOLVED IN A SIGNAL TRANSDUCTION PATHWAY THAT IS
CC ACTIVATED IN UNDER CONDITIONS OF HEAT SHOCK, OXIDATIVE STRESS OR
CC LIMITED NUTRITION. UNLIKE WISK1, IT IS NOT ACTIVATED BY CHANGES IN

```

CC THE OSMOLARITY OF THE EXTRACELLULAR ENVIRONMENT. ACTIVATES THE
CC WIS1 MAP KINASE KINASE BY PHOSPHORYLATION.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z98763; CAB1500.1; -
DR EMBL; Y07750; CAA69030.1; ALT_INIT.
DR EMBL; Y11989; CAA72718.1; -
DR EMBL; U81521; AAB39762.1; -
DR PIR; T39225; T39225.
DR HSSP; P24941; 1CKP.
DR GenDB; SPombe; SPAC9G1.02; -
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002230; Ser_thr_kinase.
DR InterPro; IPR01245; Tyr_kinase.
DR Pfam; PF00069; pkinae; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_Tkc1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS0108; PROTEIN_KINASE_ST; 1.
DR Transfase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 1037 1306 PROTEIN KINASE.
FT NP_BIND 1043 1051 ATP (BY SIMILARITY).
FT BINDING 1066 1066 ATP (BY SIMILARITY).
FT ACT_SITE 1161 1161 BY SIMILARITY.
FT FT 484 485 RL -> SP (IN REF. 3).
FT CONFLICT 1346 1346 D -> V (IN REF. 1).
SQ SEQUENCE 1401 AA; 160537 MW; 529A989D2B67F7D CRC64;

Query Match 56.3%; Score 40; DB 1; Length 1401;
Best Local Similarity 53.3%; Pred. No. 33;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 IDLASEFLSNSFL 15
DB 593 IHMAQGFSTYLSNL 607

RESULT 7
NUM_COTJA STANDARD; PRT; 72 AA.
AC P24968;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3) (Fragment).
GN MTND1 OR ND1 OR NADH1.
OS Coturnix coturnix japonica (Japanese quail).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OC NCB1_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91178819; PubMed=1706782;
RA Desjardins P., Morais R.;
RT "Nucleotide sequence and evolution of coding and noncoding regions of
RT a quail mitochondrial genome."
RT J. Mol. Evol. 32:153-161(1991).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X57246; CAA40522.1; -
DR PIR; S25422; S25422.
DR InterPro; IPR001694; Resp_NADH_dh1.
DR Pfam; PF00146; NADdh; 1.
DR PROSITE; PS00667; COMPLEX1_ND1_1; PARTIAL.
DR PROSITE; PS00668; COMPLEX1_ND1_2; PARTIAL.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
FT NON_TER 1 1
SQ SEQUENCE 72 AA; 8288 MW; F4728BBD6F650C2 CRC64;

Query Match 54.9%; Score 39; DB 1; Length 72;
Best Local Similarity 60.0%; Pred. No. 2.4;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 IDLASEFLSNSFL 15
DB 12 ITLATVLLSSFL 26

RESULT 8
NUM_METSE STANDARD; PRT; 99 AA.
AC Q47432;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN ND4L.
OS Metridium senile (Brown sea anemone) (Frilled sea anemone).
OC Mitochondrion.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actinaria;
OC Nymnathae; Metridiidae; Metridium.
OC NCB1_TaxID=6116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White morph;
RA Beagley C.T., Okimoto R., Wolstenholme D.R.;
RL Submitted (Apr-1997) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF000023; AAC04632.1; -
DR PIR; T11886; T11886.
DR InterPro; IPR003215; NADH_dh_ubiq1.
DR InterPro; IPR001133; Oxidored_4L.
DR InterPro; IPR003214; Oxidored4L.
DR Pfam; PF00420; oxidored_q2; 1.
DR ProDom; PD002107; NADH_dh_ubiq1; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 99 AA; 11035 MW; 7E8FBAAF45CCB4 CRC64;

Query Match 54.9%; Score 39; DB 1; Length 99;
Best Local Similarity 61.5%; Pred. No. 3.3;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 IASEFLPLSNSFL 15
DB 11: 11: 11:

```


FT VARIANT 64 64 S -> G (IN STRAIN PB306).
 FT VARIANT 70 70 L -> V (IN STRAIN PB306).
 FT VARIANT 317 317 G -> S (IN STRAIN PB303).
 FT VARIANT 392 409 SAPLVMTNIVFWLSVY -> QHH (IN STRAIN
 FT SEQUENCE 409 AA; 47206 MW; AFBFE452A5814F7C CRC64;
 KR114).

Query Match 54.9%; Score 39; DB 1; Length 409;
 Best Local Similarity 53.3%; Pred. No. 14;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IDLASEFLPLNSFL 15
 DB 334 LSPFLSEFLVINSML 348

RESULT 11
 SURE_CHLMU STANDARD; PRT; 263 AA.
 AC 09PRH4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Acid phosphatase sure (EC 3.1.3.2).
 GN SURE OR TC0491.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxId=83560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MOPN / N199;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA Linher O., Hickey B.K., Peterson J., Uetendack T., Berry K., Bass S.,
 RA Gwin M., Weidman J., Khouri H., Craven C., Dodson R.,
 RA Eisen W., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Bisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MOPN and Chlamydia
 pneumoniae AR39." 28:1397-1406(2000).
 RL Nucleic Acids Res.
 CC -1- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
 alcohol + phosphate.
 CC -1- COFACTOR: Magnesium (by similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
 CC -1- SIMILARITY: BELONGS TO THE SURE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AB002318; AAF39336.1; ALT_INT.
 DR HSRP; P96112; J39L.
 DR TIGR; TC0491; -.
 DR HAMAP; MF_00060; -; 1.
 DR InterPro; IPR002828; SURE.
 DR Pfam; PF01975; SURE; 1.
 DR ProDom; PD005378; SURE; 1.
 DR TIGRFAMs; TIGR00087; sure; 1.
 KM Hydrolyase; Magnesium; Complete proteome.
 FT METAL 14 14 MAGNESIUM (BY SIMILARITY).
 FT METAL 15 15 MAGNESIUM (BY SIMILARITY).
 FT METAL 47 47 MAGNESIUM (BY SIMILARITY).
 FT METAL 105 105 MAGNESIUM (BY SIMILARITY).
 FT ACT SITE 137 137 POTENTIAL.
 SQ SEQUENCE 263 AA; 31373 MW; 80E5A79A0D4F792E CRC64;
 Query Match 53.5%; Score 38; DB 1; Length 283;
 Best Local Similarity 42.9%; Pred. No. 15;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 2 DLASEFLPLNSFL 15
 DB 226 DLSEVYLAANNYI 239

RESULT 12
 TRME_LEPIN STANDARD; PRT; 456 AA.
 AC P97043;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Probable tRNA modification GTPase trme.
 GN TRME OR LA0180.
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxId=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ictero No.1 / Serogroup Icterohaemorrhagiae;
 RX MEDLINE=98332717; PubMed=966070;
 RA Takahashi Y., Akase K., Hirano H., Fukunaga M.;
 RT "Physical and genetic maps of the Leptospira interrogans serovar
 RT Icterohaemorrhagiae strain Ictero No.1 chromosome and sequencing of a
 RT 19-kb region of the genome containing the 5S rRNA gene.";
 RL Gene 215:37-45(1998).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lat;
 RX MEDLINE=22598143; PubMed=12712204;
 RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
 RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
 RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
 RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
 RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
 RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
 RA Xu J.-G., Zhao G.-P.;
 RT "Unique physiological and pathogenic features of Leptospira
 RT interrogans revealed by whole-genome sequencing.";
 RL Nature 422:888-893(2003).
 CC -1- FUNCTION: EXHIBITS A VERY HIGH INTRINSIC GTPASE HYDROLYSIS RATE.
 CC INVOLVED IN THE BIOSYNTHESIS OF THE HYPERMODIFIED NUCLEOSIDE 5-
 CC METHYLAMINOMETHYL-2-THIOURIDINE, WHICH IS FOUND IN THE WOBBLE
 CC POSITION OF SOME TRNAS (BY SIMILARITY).
 CC -1- SIMILARITY: Belongs to the era/trme family of GTP-binding
 CC proteins. Trme subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AB001721; BAA19450.1; ALT_INIT.
 DR EMBL; AB010203; BAA24373.1; ALT_INIT.
 DR EMBL; AB011208; AAN47379.1; -.
 DR PIR; T00126; T00126.
 DR HAMAP; MF_00379; -; 1.
 DR InterPro; IPR005289; GTP-binding_dom.
 DR InterPro; IPR002917; MMR_HSR1.
 DR InterPro; IPR005225; Small_GTP.
 DR InterPro; IPR004520; ThdF.
 DR Pfam; PF01926; MMR_HSR1; 1.
 DR PRINTS; PR00449; RASTRNSFRNG.
 DR TIGRFAMs; TIGR00650; MG442; 1.
 DR TIGRFAMs; TIGR00231; smg11_GTP; 1.
 DR TIGRFAMs; TIGR00450; thdF; 1.
 DR tRNA processing; GTP-binding; Complete proteome.
 NP_BIND 227 234 GTP (POTENTIAL).
 NP_BIND 274 278 GTP (POTENTIAL).

Query Match
Best Local Similarity 53.5%; Score 38; DB 1; Length 456;
Best Local Similarity 46.7%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 IDLASEPLFLNSPL 15
DB 28 LTSSSFLSKXKFL 42

RESULT 13
YQ37_SCHPO STANDARD; PRT; 735 AA.

AC Q10432;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C188.07 in chromosome III.
GN SPC188.07.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
NCBI_Taxid=4896;
RX MEDLINE=21848401; PubMed=11859360;
RC STRAIN=972;
RA Wood V., Gilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Brooks J., Peat N., Hayles J., Baker S., Besham D., Bowman S.,
RA Sgouras K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornbly T., Howarth S., Huckle E.J., Hunt S., Jagsel K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy J., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkhardt G., Aert R., Robben J., Grynoprez B.,
RA Welljens I., Vanstreels B., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langner I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AL049662; CAB41227.1; -
DR PIR: T41187; T41187 -
DR GeneDB_Spomb; SPC188.07; -
KW Hypothetical protein.
SQ SEQUENCE 735 AA; 83300 MW; 14193682700A2634 CRC64;

Query Match
Best Local Similarity 53.5%; Score 38; DB 1; Length 735;
Best Local Similarity 77.8%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 FIFLNSPL 15
DB 245 FLFISNPL 253

RESULT 14
AF6_HUMAN STANDARD; PRT; 1816 AA.

AC P55196; O75087; O75088; O75089; Q9NU92;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE AF-6 protein.
GN MLT4 OR AF6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_Taxid=9606;
RX MEDLINE=94061833; PubMed=8242616;
RA Prasad R., Gu Y., Alder H., Nakamura T., Kanamori O., Saito H.,
RA Huebner K., Gale R.P., Nowell P.C., Kuriyama K., Miyazaki Y.,
RA Croce C.M., Canani E.,
RT "Cloning of the ALU-1 fusion partner, the AF-6 gene, involved in
RT acute myeloid leukemias with the t(6;11) chromosome translocation.";
RL Cancer Res. 53:5624-5628(1993).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL AB011399; BA332484.1; -
DR EMBL AB011399; BA332483.1; -
DR EMBL AB011399; BA332485.1; -
DR EMBL U02478; AAC50059.1; -
DR EMBL AL049698; CAB76850.1; -
DR HSSP; 012923; 3PDZ.
DR Genew; HGNC:7137; MLT4.
DR MIM; 159589; -
DR GO; GO:0005911; C:intercellular junction; TMS.
DR GO; GO:0006022; F:protein C-terminus binding activity; TMS.
DR GO; GO:0007155; P:cell adhesion; TMS.
DR GO; GO:0007267; P:cell-cell signaling; TMS.
DR GO; GO:0007048; P:oncogenesis; TMS.
DR InterPro; IPR002710; DIL.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR000159; RA_domain.
DR Pfam; PF01843; DIL; 1.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00595; PDZ; 1.
DR Pfam; PF00788; RA; 2.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00228; PDZ; 1.
DR SMART; SM00314; RA; 2.
DR PROSITE; PS50106; PDZ; 1.
DR PROSITE; PS50200; RA; 2.
KM Chromosomal translocation; Proto-oncogene; Alternative splicing;
Repeat.
FT DOMAIN 39 133 RAS-ASSOCIATING 1.
FT DOMAIN 245 347 RAS-ASSOCIATING 2.
FT DOMAIN 425 481 FHA.
FT DOMAIN 804 910 DILUTE.
FT DOMAIN 991 1077 PDZ.
FT DOMAIN 162 174 GUU/LYS-RICH.
FT DOMAIN 1349 1356 POLY-PRO.
FT DOMAIN 1371 1376 ASP/GU-RICH (ACIDIC).
FT DOMAIN 1561 1571 ML FUSION POINT.
FT SITE 26 26 (IN AN ACUTE MYELOID LEUKEMIA PATIENT).
FT VARSPLIC 1588 1611 LODDERROCOLEEMKREAPRA -> VKGVLMCPSPV
FT FT /FTId=VSP_000217.
FT FT /FTId=VSP_000219.
FT FT /FTId=VSP_000218.
FT VARSPLIC 1612 1816 Missing (in isoform 1).
FT FT /FTId=VSP_000218.
FT VARSPLIC 1666 1743 LCRPPLPRDYEPSPSPAPGAPPPPPORNASYLKTOVLSPD
FT FT SLFTAKFVAVNEEEDCDLAGODYSTRKSHGDL ->
FT FT PMSYPGSTGAAGAHDCADKAKREKSKQDADSPGSGAPE
FT FT NLTFRORLPSQGDGVSKVAKSRKLTLENEMLNK (in
FT FT isoform 3).
FT FT /FTId=VSP_000219.
FT FT /FTId=VSP_000218.
FT VARSPLIC 1744 1816 Missing (in isoform 3).
FT FT /FTId=VSP_000220.
FT FT G -> V (IN REF. 1).
FT FT P -> PGRNHPAYVHYHYE (IN REF. 3).
FT FT CONFLICT 373 373 D -> DSHFD (IN REF. 3).
FT FT CONFLICT 391 391 D -> DV (IN REF. 1).
FT FT CONFLICT 744 744 R -> P (IN REF. 1).
FT FT CONFLICT 1031 1031
FT FT CONFLICT 1408 1408
FT FT CONFLICT 1816 AA; 205604 MW; EB1FE7F04879CE8F CR64;
SQ SEQUENCE
Query Match 53.5%; Score 38; DB 1; Length 1816;
Beet Local Similarity 72.7%; Pred. No. 1e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 4 ASEPLLSNSF 14
DB 1760 ASDGVFLNSF 1770
RESULT 15
YCBT_BACSU STANDARD; PRT; 245 AA.
AC P42252; OJ1468;

```

```

DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ycbT.
GN YCBT.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC NCBI_Taxid=1423;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=95219079; PubMed=7704254;
RT Ogawa K.-I., Akagawa E., Nakamura K., Yamane K.;
RT "Determination of a 21548 bp nucleotide sequence around the 24
RT degrees region of the Bacillus subtilis chromosome.";
RT Microbiology 141:269-275 (1995).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RT Kunze F., Ogatawara N., Moszer I., Albertini A.M., Alloni G.,
RT Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RT Bourlles R., Boursier L., Brans A., Braun M., Brignelli S.C., Bron S.,
RT Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RT Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RT Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RT Eutlan K.D., Etrington J., Fabre C., Ferrari E., Foulger D.,
RT Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RT Ghim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
RT Giuseppe G., Guy B.J., Haga K., Halech U., Harwood C.R., Henaut A.,
RT Hilbert H., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RT Joris B., Karamata D., Kasahara Y., Klaerr-Bianchini N., Klein C.,
RT Kobayashi Y., Koelster P., Koningsstein G., Krogh S., Kumano M.,
RT Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RT Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RT Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RT Noone D., O'Reilly M., Ogawa K., Ogiwara A., Ondega B., Park S.H.,
RT Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RT Priescan E., Pujic P., Purnelle E., Rapoport G., Ray M., Reynolds S.,
RT Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RT Sato T., Scanlan E., Schleich S., Schroeder R., Scofield F.,
RT Sekiguchi J., Sekowska A., Serró S.J., Serró P., Shin B.S., Soldo B.,
RT Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RT Takeuchi M., Tamakoshi A., Tanaka T., Terpeira P., Tognoni A.,
RT Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vasarotti A.,
RT Viari A., Wambut R., Wedler E., Wedler H., Weltzenegger T.,
RT Wintere P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RT Yoshida K., Yoshikawa H.F., Zumschein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RT Nature 390:249-256 (1997).
RN (3)
RP SEQUENCE OF 190-245 FROM N.A.
RC STRAIN=168;
RX MEDLINE=92339527; PubMed=1353026;
RT Awade A., Cleuziat P., Gonzalez T., Robert-Baudouy J.;
RT "Characterization of the pep gene encoding the pyrrolidone carboxyl
RT peptidase of Bacillus subtilis."
RT FEBS Lett. 305:67-73 (1992).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: BELONGS TO THE TARC FAMILY.
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 20.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; D30808; BAA06484.1; ALT_FRAME.

```

DR EMBL; Z99105; CAB12058.1; -.
DR EMBL; X66034; -; NOT_ANNOTATED_CDS.
DR PIR; D69754; D69754.
DR Subtilist; BG11175; YcbT.
DR InterPro; IPR002033; Translocase.
DR Pfam; PF00902; Talc; 1.
DR PROSITE; PS01218; TATC; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 60 80 POTENTIAL.
FT TRANSMEM 110 130 POTENTIAL.
FT TRANSMEM 153 173 POTENTIAL.
FT TRANSMEM 191 211 POTENTIAL.
FT TRANSMEM 212 232 POTENTIAL.
SQ SEQUENCE 245 AA; 28114 MW; D5066B68056EC231 CRC64;

Query Match 52.1%; Score 37; DB 1; Length 245;
Best local Similarity 53.3%; Pred. No. 19;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IDLASEFLPLNSFL 15
Db 21 VTLAFLPLITRAFL 35

Search completed: November 26, 2003, 12:31:20
Job time : 8.22892 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:25:40 ; Search time 13.9157 Seconds
(Without alignments)
103.662 Million cell updates/sec

Title: US-09-230-111C-15

Perfect score: 71

Sequence: 1 IDLASEFLFLSNSFL 15

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR 76:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	42	59.2	98	2	SS3467	hypothetical prote
2	41	57.7	106	2	SS1624	FRAC-1 protein -
3	41	57.7	106	2	SS1625	FRAC-2 protein -
4	41	57.7	376	2	I49735	hepatocyte nuclear
5	41	57.7	444	2	T09474	forkhead protein P
6	40	56.3	554	2	S71751	dolichyl-phosphate
7	40	56.3	1401	2	T39225	MAP kinase kinase
8	39	54.9	72	2	S25422	NADH2 dehydrogenas
9	39	54.9	99	2	T11886	NADH2 dehydrogenas
10	39	54.9	164	2	T12408	NADH2 dehydrogenas
11	39	54.9	409	2	S26033	NADH2 dehydrogenas
12	39	54.9	819	2	T40798	hypothetical prote
13	39	54.9	1170	2	T27664	hypothetical prote
14	39	54.9	3724	2	T18427	hypothetical prote
15	38	53.5	134	2	D95062	hypothetical prote
16	38	53.5	291	2	A81696	immunity protein B
17	38	53.5	435	2	T00126	stationary-phase s
18	38	53.5	548	2	AD1216	hypothetical prote
19	38	53.5	548	2	A11569	ABC transporters,
20	38	53.5	735	2	T41187	ABC transporters,
21	38	53.5	1110	2	T29327	hypothetical prote
22	37	52.1	111	2	D97792	hypothetical prote
23	37	52.1	201	2	C96813	hypothetical prote
24	37	52.1	245	2	D69754	conserved hypotet
25	37	52.1	309	2	E71019	NADH2 dehydrogenas
26	37	52.1	309	2	H75113	NADH2 dehydrogenas
27	37	52.1	324	2	S10187	NADH2 dehydrogenas
28	37	52.1	411	1	SYBYCS	carbamoyl-phosphat
29	37	52.1	590	2	C81316	probable sugar epi

30	37	52.1	755	2	C90439	cation transportin
31	37	52.1	1178	2	S78475	mannosylphosphoryl
32	37	52.1	1210	2	D88013	protein K10B4.1 (1
33	37	52.1	1436	2	A99115	putative U5 snRNP-
34	36.5	51.4	705	2	T49461	probable cro1 prot
35	36	50.7	164	1	ZPECL	signal peptidase 1
36	36	50.7	164	2	P90632	polipoprotein sig
37	36	50.7	164	2	E85483	signal peptidase 1
38	36	50.7	166	2	AB0508	lipoprotein signal
39	36	50.7	172	2	A64168	hypothetical prote
40	36	50.7	249	2	C90526	conserved hypotet
41	36	50.7	251	2	G89928	conserved hypotet
42	36	50.7	293	2	G69854	conserved hypotet
43	36	50.7	294	2	E69817	hypothetical prote
44	36	50.7	342	2	T29192	hypothetical prote
45	36	50.7	346	1	B64976	galactitol-1-phosp

ALIGNMENTS

RESULT 1
S53467
hypothetical protein YAR053w - yeast (Saccharomyces cerevisiae)
C/Species: Saccharomyces cerevisiae
C/Date: 05-May-1995 #sequence_revision 01-Sep-1995 #ext_change 19-Apr-2002
C/Accession: S53467
R/Busey, H.; Keng, T.; Storms, R.K.; Vo, D.; Zhong, W.; Fortin, N.; Barton, A.B.; Kabe
submitted to the EMBL Data Library, February 1994
A/Description: Sequencing of chromosome I of Saccharomyces cerevisiae: analysis of the
A/Reference number: S53458
A/Accession: S53467
A/Molecule type: DNA
A/Residues: 1-98 <BUS>
A/Cross-references: EMBL:L28920; NID:G1616966; PID:G456147; GSPDB:GN00001; MIPS:YAR053w
C/Genetic: A/Genes: MIPS:YAR053w
A/Cross-references: SGD:S0000085
A/Map position: IR

Query Match
Best Local Similarity 59.2%; Score 42; DB 2; Length 98;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 IDLASEFLFLSNSF 14
Db 75 VMLPCSFIRISNSF 88

RESULT 2

S51624
FRAC-1 protein - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #ext_change 07-May-1999
C/Accession: S51624
R/Pierrou, S.; Hellqvist, M.; Samuelsson, L.; Enerbaeck, S.; Carlsson, P.
EMBO J. 13, 5003-5012, 1994
A/Title: Cloning and characterization of seven human forkhead proteins: binding site sp.
A/Reference number: S51624; MUID:95045392; PMID:7957066
A/Accession: S51624
A/Status: preliminary; nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-106 <PIE>
A/Cross-references: EMBL:U13219
C/Genetic: A/Genes: GDB:FKHL5; FRAC1
A/Cross-references: GDB:450216; OMIM:601089
A/Map position: 16q24-16q24
C/Supplementary: unassigned fork head proteins; fork head DNA-binding domain homology
F/6-97/Domain: fork head DNA-binding domain homology <FHD>

Query Match
Best Local Similarity 57.7%; Score 41; DB 2; Length 106;
Matches 64.3%; Pred. No. 2.6;

Matches: 9; Conservative: 1; Mismatches: 4; Indels: 0; Gaps: 0;

QY 1 IDLASEFLSNSF 14
|||:|
DB 82 IDPASEFMEEGSF 95

RESULT 3
S51625
FRRAC-2 protein - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 07-May-1999
C/Accession: S51625
R/Pierrou, S.; Hellqvist, M.; Samuelsson, L.; Eneback, S.; Carlsson, P.
EMBO J. 13, 5002-5012, 1994
A/Title: Cloning and characterization of seven human forkhead proteins: binding site spe
A/Reference number: S51624; MWID:95045392; PMID:7957066
A/Accession: S51625
A/Status: preliminary; nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-106 <PIE>
A/Cross-references: EMBL:U13220
C/Genetics:
A/Gene: GDB:FKHL6; FRRAC2
A/Cross-references: GDB:450218
A/Map position: 16q24-16q24
A/Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
F:6-97//Domain: fork head DNA-binding domain homology <FHD>

Query Match 57.7%; Score 41; DB 2; Length 106;
Best Local Similarity 64.3%; Pred. No. 2.6;
Matches: 9; Conservative: 1; Mismatches: 4; Indels: 0; Gaps: 0;

QY 1 IDLASEFLSNSF 14
|||:|
DB 82 IDPASEFMEEGSF 95

RESULT 4
149735
hepatocyte nuclear factor 3 forkhead homolog 8 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 24-Sep-1999
C/Accession: I49735
R/Clevidence, D.E.; Overdier, D.G.; Peterson, R.S.; Porcella, A.; Ye, H.; Paulson, K.E.;
Dev. Biol. 166, 193-209, 1994
A/Title: Members of the HNF-3/forkhead family of transcription factors exhibit distinct
A/Reference number: I49735; MWID:95046902; PMID:7958446
A/Accession: I49735
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-376 <RES>
A/Cross-references: GB:LI35949; NID:G762833; PIDN:AAA64885.1; PID:G575512
C/Genetics:
A/Gene: HFN-8
C/Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
F:47-138//Domain: fork head DNA-binding domain homology <FHD>

Query Match 57.7%; Score 41; DB 2; Length 376;
Best Local Similarity 64.3%; Pred. No. 10;
Matches: 9; Conservative: 1; Mismatches: 4; Indels: 0; Gaps: 0;

QY 1 IDLASEFLSNSF 14
|||:|
DB 123 IDPASEFMEEGSF 136

RESULT 5
T09474
forkhead protein FRRAC-2 - human
C/Species: Homo sapiens (man)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000
C/Accession: T09474

R/Hellqvist, M.; Mahlapuu, M.; Blixt, C.; Eneback, S.; Carlsson, P.
J. Biol. Chem. 273, 23335-23343, 1998
A/Title: The human forkhead protein FRRAC-2 contains two functionally redundant activat
A/Reference number: Z16682; MWID:98389768; PMID:9722567
A/Accession: T09474
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-444 <HEL>
A/Cross-references: EMBL:U13220; NID:G3425849; PID:G3425850
A/Experimental source: lung
C/Function:
A/Description: may function as transcription activator
C/Superfamily: transcription factor HNF-3; fork head DNA-binding domain homology
F:100-191//Domain: fork head DNA-binding domain homology <FHD>

Query Match 57.7%; Score 41; DB 2; Length 444;
Best Local Similarity 64.3%; Pred. No. 12;
Matches: 9; Conservative: 1; Mismatches: 4; Indels: 0; Gaps: 0;

QY 1 IDLASEFLSNSF 14
|||:|
DB 176 IDPASEFMEEGSF 189

RESULT 6
S71751
dolichyl-phosphate-mannose-glycolipid alpha-mannosyltransferase (EC 2.4.1.130) FIG-B - h
N/Alternate names: membrane protein FIG-B; phosphatidyl-inositol glycolipid biosynthesi
C/Species: Homo sapiens (man)
C/Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 03-Jun-2002
C/Accession: S71751
R/Rakhaashi, M.; Inoue, N.; Ohishi, K.; Maeda, Y.; Nakamura, N.; Endo, Y.; Fujita, T.;
EMBO J. 15, 4254-4261, 1996
A/Title: FIG-B, a membrane protein of the endoplasmic reticulum with a large luminal do
A/Reference number: S71751; MWID:97015126; PMID:8861954
A/Accession: S71751
A/Molecule type: mRNA
A/Residues: 1-554 <TAK>
A/Cross-references: EMBL:D42138; NID:G1552168; PIDN:BA07709.1; PID:G1552169
A/Experimental source: cell line P39
A/Note: Part of the genomic DNA was also sequenced
C/Genetics:
A/Gene: GDB:FIGB
A/Cross-references: GDB:9956843; OMIM:604122
A/Map position: 15q21-15q22
C/Function:
A/Description: involved in surface protein binding to the membrane via glycosyl-phospha
A/Pathway: GPI-anchor biosynthesis
C/Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein
F:60-77//Domain: transmembrane #status predicted <TMM>

Query Match 56.3%; Score 40; DB 2; Length 554;
Best Local Similarity 75.0%; Pred. No. 23;
Matches: 9; Conservative: 0; Mismatches: 3; Indels: 0; Gaps: 0;

QY 4 ASEFLPLSNSF 15
|:|:|:|
DB 388 ALSFLPLSNLF 399

RESULT 7
T39225
MAP kinase kinase kinase - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
C/Accession: T39225
R/Churcher, C.M.; Gentles, S.; Bartell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997
A/Reference number: Z21837
A/Accession: T39225
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1401 <CHU>

A/Cross-references: EMBL:Z98763; PIDN:CAB1500.1; GSPDB:GN00066; SPDB:SPAC9G1.02
 A/Experimental source: strain 97zh-; cosmid c9G1
 C/Genetics:
 A/Gene: SPDB:SPAC9G1.02
 A/Map position: 1
 C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

Query Match 56.3%; Score 40; DB 2; Length 140;
 Best Local Similarity 53.3%; Pred. No. 63;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IDLASEFLPLNSFL 15
 |||:|||||
 Db 593 IHMACPFYLSNSL 607

RESULT 8
 S25422
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 1 - Japanese quail mitochondrion (fr
 C/Species: mitochondrion Coturnix coturnix japonica (Japanese quail)
 C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
 C/Accession: S25422
 R/DeeJardine, P.; Morals, R.
 J. Mol. Evol. 32, 153-161, 1991
 A/Title: Nucleotide sequence and evolution of coding and noncoding regions of a quail m
 A/Reference number: S25422; MUID:91178819; PMID:1706782
 A/Accession: S25422
 A/Molecule type: DNA
 A/Residues: 1-72 <DES>
 A/Cross-references: EMBL:X57246; NID:g12849; PIDN:CAA40522.1; PID:g12850
 C/Genetics:
 A/Gene: ND1
 A/Genome: mitochondrion
 A/Genetic code: SGC1
 C/Superfamily: NADH dehydrogenase (ubiquinone) chain 1
 C/Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 54.9%; Score 39; DB 2; Length 72;
 Best Local Similarity 60.0%; Pred. No. 4.1;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IDLASEFLPLNSFL 15
 |||:|||||
 Db 12 ITLATKULLSSFL 26

RESULT 9
 T11886
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - sea anemone (Metridium senille)
 C/Species: mitochondrion Metridium senille (brown sea anemone, filled sea anemone)
 C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
 C/Accession: T11886
 R/Beegley, C.T.; Okimoto, R.; Wolstenholme, D.R.
 submitted to the EMBL Data Library, April 1997
 A/Reference number: Z17371
 A/Accession: T11886
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Residues: 1-25 <OK2>
 A/Molecule type: DNA
 A/Residues: 1-99 <BEA>
 A/Cross-references: EMBL:AF000023; NID:g2920983; PID:g2920987; PIDN:AAC04632.1
 A/Experimental source: white color morph
 C/Genetics:
 A/Gene: ND4L
 A/Genome: mitochondrion
 A/Note: ND4L
 C/Superfamily: NADH dehydrogenase (ubiquinone) chain 4L
 C/Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 54.9%; Score 39; DB 2; Length 99;
 Best Local Similarity 61.5%; Pred. No. 5.7;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LASEFLPLNSFL 15
 ||:|||||
 Db

Db 39 LAASEFLPLNSML 51

RESULT 10
 T12408
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Heterorhabditis hepialus mitochc
 C/Species: mitochondrion Heterorhabditis hepialus
 C/Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 03-Jun-2002
 C/Accession: T12408
 R/Liu, J.; Berry, R.; Blouin, M.
 submitted to the EMBL Data Library, May 1998
 A/Description: Molecular differentiation and phylogeny of entomopathogenic nematodes.
 A/Reference number: Z17506
 A/Accession: T12408
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-164 <LIU>
 A/Cross-references: EMBL:AF066880; NID:g3192836; PID:g3192837; PIDN:AAC19101.1
 A/Experimental source: strain BOD
 C/Genetics:
 A/Gene: ND4
 A/Genome: mitochondrion
 A/Note: ND4
 C/Superfamily: NADH dehydrogenase (ubiquinone) chain 4
 C/Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 54.9%; Score 39; DB 2; Length 164;
 Best Local Similarity 46.7%; Pred. No. 9.8;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IDLASEFLPLNSFL 15
 |||:|||||
 Db 130 LSFISEFLVINSGL 144

RESULT 11
 S26033
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Caenorhabditis elegans mitochond
 C/Species: mitochondrion Caenorhabditis elegans
 C/Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 03-Jun-2002
 C/Accession: S26033; S25806
 R/Okimoto, R.; Macfarlane, J.L.; Clary, D.O.; Wolstenholme, D.R.
 Genetics 130, 471-498, 1992
 A/Title: The mitochondrial genomes of two nematodes, Caenorhabditis elegans and Ascaris
 A/Reference number: S26014; MUID:92201635; PMID:1551572
 A/Accession: S26033
 A/Molecule type: DNA
 A/Residues: 1-409 <OKI>
 A/Cross-references: EMBL:X54252; NID:g13988; PIDN:CAA38158.1; PID:g559502
 A/Note: the authors translated the initiation codon TTG for residue 1 as Leu
 R/Okimoto, R.; Macfarlane, J.L.; Wolstenholme, D.R.
 Nucleic Acids Res. 18, 6113-6118, 1990
 A/Title: Evidence for the frequent use of TTG as the translation initiation codon of mt
 A/Reference number: S13139; MUID:91045077; PMID:2235493
 A/Accession: S25806
 A/Molecule type: DNA
 A/Residues: 1-25 <OK2>
 A/Cross-references: EMBL:X54252
 A/Note: the authors translated the initiation codon TTG for residue 1 as Leu
 C/Genetics:
 A/Gene: ND4
 A/Genome: mitochondrion
 A/Genetic code: SGC4
 A/Start codon: TGG
 C/Superfamily: NADH dehydrogenase (ubiquinone) chain 4
 C/Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 54.9%; Score 39; DB 2; Length 409;
 Best Local Similarity 53.3%; Pred. No. 26;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IDLASEFLPLNSFL 15
 |||:|||||
 Db 334 LSFISEFLVINSML 348

RESULT 12

hypothetical protein SPBP87.04 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000

C/Accession: T40798

R:Beck, A.; Reinhardt, R.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, October 1998

A:Reference number: Z21949

A/Accession: T40798

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Accession: T40798

A/Molecule type: DNA

A/Residues: 1-819 <BEC>

A/Cross-references: EMBL:AL032684; PIDN:CA21789.1; GSPDB:GN00067; SPDB:SPBP87.04

A/Experimental source: strain 972h-; clone pl p887

C/Genetics:

A:Gene: SPDB:SPBP87.04

A:Map position: 2

A:introns: 24/3

C:Superfamily: Schizosaccharomyces hypothetical protein SPBP87.04

Query Match 54.9%; Score 39; DB 2; Length 819;

Best Local Similarity 53.8%; Pred. No. 54;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 DLASEFLFINSFL 14

DB 756 DKSEFLIQNSY 768

RESULT 13

hypothetical protein ZC412.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000

C/Accession: T27564

R:Ainscough, R.

submitted to the EMBL Data Library, August 1996

A:Reference number: Z20387

A/Accession: T27564

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1170 <WIL>

A/Cross-references: EMBL:Z78067; PIDN:CA80153.1; GSPDB:GN00023; CESP:ZC412.2

A/Experimental source: clone ZC412

C/Genetics:

A:Gene: CESP:ZC412.2

A:Map position: 5

A:introns: 65/2; 146/2; 222/1; 264/1; 304/3; 345/3; 435/3; 500/2; 834/3; 913/2; 977/3; 1

C:Superfamily: membrane-bound guanylate cyclase; guanylate cyclase catalytic domain hom

Query Match 54.9%; Score 39; DB 2; Length 1170;

Best Local Similarity 50.0%; Pred. No. 79;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 DLASEFLFINSFL 15

DB 703 DISNGLFTINSFL 716

RESULT 14

hypothetical protein C0335c - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C/Accession: T18427

R:Lawson, D.; Bowman, S.; Barrell, B.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z18935

A/Accession: T18427

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-3724 <IAM>

A/Cross-references: EMBL:Z98547; NID:e1325376; PID:e1325379; PIDN:CA81104.1

C/Genetics:

A:introns: 307/1; 1545/2

A>Note: C0335c

Query Match 54.9%; Score 39; DB 2; Length 3724;

Best Local Similarity 46.7%; Pred. No. 2.7e+02;

Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 IDLASEFLFINSFL 15

DB 3130 VIDTEFLTYRYFL 3144

RESULT 15

immunity protein B1PL [imported] - Streptococcus pneumoniae (strain TIGR4)

C:Species: Streptococcus pneumoniae

C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001

C/Accession: D95062

R:Teitelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid

on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,

nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A>Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A/Accession: D95062

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-134 <KUR>

A/Cross-references: GB:AE005672; PIDN:AAK74693.1; PID:G14972010; GSPDB:GN00164; TIGR:SP

A/Experimental source: strain TIGR4

C/Genetics:

A:Gene: SP0536

Query Match 53.5%; Score 38; DB 2; Length 134;

Best Local Similarity 61.5%; Pred. No. 12;

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 IASEFLFINSFL 15

DB 24 LVSTYFLFSRQFL 36

Search completed: November 26, 2003, 12:36:06
Job time: 15.9157 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:25:10 ; Search time 32.7108 Seconds
(without alignments)
118.334 Million cell updates/sec

Title: US-09-230-111c-15
Perfect score: 71
Sequence: 1 IDLASEPLFLSNSFL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacterioplasmid:*
17: sp_archaea:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	64.8	53	16	Q8F570 leptospira
2	46	64.8	82	17	Q8HJK3 thermoplasma
3	45	63.4	682	10	Q9SR73 arabidopsis
4	43	60.6	82	17	Q979M6 thermoplasma
5	41	57.7	267	8	Q8W147 gyrocarpus
6	41	57.7	295	13	Q9VVP9 gallus galli
7	41	57.7	360	8	Q9GHRV hermannia n
8	41	57.7	372	13	Q9W707 xenopus lae
9	41	57.7	373	13	Q9W706 xenopus lae
10	41	57.7	446	11	Q54743 mus musculus
11	40	56.3	164	8	Q63861 heterorhabd
12	40	56.3	164	8	Q63862 heterorhabd
13	40	56.3	164	8	Q63860 heterorhabd
14	40	56.3	186	2	Q9SE50 streptococc
15	40	56.3	245	3	Q9HFU7 pneumococc
16	40	56.3	252	2	Q54977 streptococc

17	40	56.3	280	2	Q85731 streptococc
18	40	56.3	280	16	Q9FB04 streptococc
19	40	56.3	280	16	Q8N271 streptococc
20	40	56.3	280	16	Q8K5P1 streptococc
21	40	56.3	421	5	Q8I1C3 plasmidium
22	40	56.3	502	10	Q9AVQ2 solanum tub
23	40	56.3	554	4	Q8WVW7 homo sapien
24	40	56.3	554	4	Q92521 homo sapien
25	40	56.3	848	17	Q8P247 methanobact
26	39	54.9	164	8	Q63932 heterorhabd
27	39	54.9	164	8	Q63859 heterorhabd
28	39	54.9	164	8	Q63858 heterorhabd
29	39	54.9	164	8	Q63933 heterorhabd
30	39	54.9	164	8	Q63982 heterorhabd
31	39	54.9	166	2	Q9RF47 klebsiella
32	39	54.9	217	11	Q9C5S4 mus musculus
33	39	54.9	324	8	Q8SEK3 coturnix co
34	39	54.9	819	3	Q94254 schizosacch
35	39	54.9	852	13	Q9PTES leptospira
36	39	54.9	1170	5	Q23310 caenorhabd
37	39	54.9	1847	5	Q81DM0 plasmidium
38	39	54.9	3724	5	Q77320 plasmidium
39	38.5	54.2	38	16	Q8F7N4 leptospira
40	38	53.5	95	2	Q8RM64 bacteroides
41	38	53.5	127	11	Q8BNG5 mus musculus
42	38	53.5	134	16	Q9ETK8 streptococc
43	38	53.5	146	2	Q9AKA8 rickettsia
44	38	53.5	377	13	Q8AW25 brachydanio
45	38	53.5	403	5	Q9VP08 drosophila

ALIGNMENTS

RESULT 1	Q8F570	PRELIMINARY;	PRT;	53 AA.
ID	Q8F570	PRELIMINARY;	PRT;	53 AA.
AC	Q8F570	01-MAR-2003 (TREMBLrel. 23, Created)		
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DE	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
GN	L1817.			
OS	Leptospira interrogans.			
OC	Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.			
OX	NCBI_TaxID=173;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	STRAIN-56601 / Serogroup Icterohaemorrhagiae / Serovar lai;			
RT	Ren S.;			
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AE01357; AAN49015.1; -			
KW	Hypothetical protein; Complete proteome.			
SQ	SEQUENCE 53 AA; 6215 MW; BFF5A0A81348719 CRC64;			
Query Match	Best Local Similarity	64.8%;	Score 46;	DB 16; Length 53;
Matches	10; Conservative	83.3%;	Pred. No. 0.61;	
			Mismatches	2; Indels
				Gaps
QY	2	DLASEPLFLSNS	13	
DB	18	DLASEPLFLSNS	29	
RESULT 2	Q8HJK3	PRELIMINARY;	PRT;	82 AA.
ID	Q8HJK3	01-MAR-2001 (TREMBLrel. 16, Created)		
AC	Q8HJK3	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)			
DE	NADH dehydrogenase, chain J related protein.			
GN	TA0964.			

```

OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Granl W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mews H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermophilic scavenger Thermoplasma
RT acidophilum."
RL Nature 407:508-513 (2000).
DR EMBL; AF45066; CAC12093.1; -.
KW Complete proteome.
SQ SEQUENCE 82 AA; 9081 MW; 04A016397953670D CRC64;

Query Match 64.8%; Score 46; DB 17; Length 82;
Best Local Similarity 66.7%; Pred. No. 0.93;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IDLASEFLPLNSFL 15
Db 39 IILASEFLPLGASFI 53

RESULT 3
Q9SRT3 PRELIMINARY; PRT; 682 AA.
AC Q9SRT3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE F2103.4 protein (Hypothetical protein).
GN F2103.4 OR AT3G07330.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA Rensing C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Frazer C.M.;
RT "Arabidopsis thaliana chromosome III BAC F2103 genomic sequence."
RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Yamada K., Banb U., Chang C.H., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Totium M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shin P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene At3g07330 (GI:15231448)."
RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Huan V.W.,
RA Lee J.M., Quach H.L., Tang C., Totium M., Wallender E.K., Wong C.,
RA Wu H.C., Yu G., Yuan S., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shin P., Southwick A., Tripp M.G., Wu T., Shinzaki K., Davis R.W.,
RA Ecker J.R., Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones."
RT Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC009853; AA02144.1; -.
DR EMBL; AY070454; AA19857.1; -.

```

```

DR EMBL; AY142677; AA113215.1; -.
DR InterPro; IPR001173; Glyco_trans_2.
DR Pfam; PF00535; Glycos_transf_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 682 AA; 78371 MW; F6D0494C506E57B9 CRC64;

Query Match 63.4%; Score 45; DB 10; Length 682;
Best Local Similarity 69.2%; Pred. No. 11;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 IASEFLPLNSFL 15
Db 70 IASEFLPLGASFI 82

RESULT 4
Q979M6 PRELIMINARY; PRT; 82 AA.
AC Q979M6;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE NADH dehydrogenase I chain G.
GN TV1134 OR TVG1162541.
OS Thermoplasma volcanium.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=50339;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=GSSI / DSM 4299 / JCM 9571;
RX MEDLINE=20570466; PubMed=11121031;
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
RA Kawashima O-hya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,
RA Nunoshima T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;
RT "Archaeal adaptation to higher temperatures revealed by genomic
RT sequence of Thermoplasma volcanium."
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262 (2000).
DR EMBL; AP000995; BAB60276.1; -.
KW Complete proteome.
SQ SEQUENCE 82 AA; 9039 MW; FA97DE7507A29966 CRC64;

Query Match 60.6%; Score 43; DB 17; Length 82;
Best Local Similarity 69.2%; Pred. No. 3.3;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 IASEFLPLNSFL 15
Db 41 IASEFLPLGASFI 53

RESULT 5
Q8W147 PRELIMINARY; PRT; 267 AA.
AC Q8W147;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Maturase K (Intin maturase) (Fragment).
GN MATK.
OS Gyrocarpus americanus.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Hernandiaceae; Gyrocarpus.
OX NCBI_TaxID=63807;
RN (1)
RP SEQUENCE FROM N.A.
RA Chatterou L.M., Savolainen V., Powell M., Bakker F.T., Zanis M.,
RA Chase M.W.;
RT "Slowly and rapidly evolving genes and phylogeny reconstruction in the
RT angiosperms: an empirical assessment."
RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II

```

```

CC INTRONS (BY SIMILARITY).
CC -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC MITOCHONDRIAL INTRONS.
CC EMBL: AF465295; AAL69409.1; -.
CC DR InterPro: IPR000442; Intron_mature2.
CC DR InterPro: IPR002866; MatK_N.
CC DR Pfam: PF01348; Intron_mature2; 1.
CC DR Pfam: PF01824; MatK_N; 1.
CC MRNA processing; Chloroplast.
CC FT NON_TER 1
CC FT 267
CC SQ SEQUENCE 267 AA; 31607 MW; 5162C649702507C1 CRC64;

Query Match 57.7%; Score 41; DB 8; Length 267;
Best Local Similarity 88.9%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 7 FLFLSNSFL 15
Db 65 FLFLSNSFLV 73

RESULT 6
O9VPV9 PRELIMINARY; PRT; 295 AA.
AC O9VPV9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Forward transcription factor (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-99387980; PubMed10457021;
RA Funayama N., Sato Y., Matsumoto K., Ogura T., Takahashi Y.;
RT "Coelom formation: binary decision of the lateral plate mesoderm is
RT controlled by the ectoderm.";
RL Development 126:4129-4138(1999).
DR EMBL: AB028627; BAA84095.1; -.
DR HSP; O63245; 2HFH.
DR InterPro: IPR001766; TF_Fork_head.
DR Pfam: PF00250; Fork_head; 1.
DR ProDom: PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS50039; FORK_HEAD_3; 1.
FT NON_TER 1
FT 295
FT SQ SEQUENCE 295 AA; 31829 MW; 7A05EE300CDA2839 CRC64;

Query Match 57.7%; Score 41; DB 13; Length 295;
Best Local Similarity 64.3%; Pred. No. 27;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IDLASEFLSNSF 14
Db 52 IDLASEFLMEEGSF 65

RESULT 7
O9GHV9 PRELIMINARY; PRT; 360 AA.
AC O9GHV9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Intron maturase (Maturase K) (Fragment).
GN MATK.

```

```

OS Hernandia nymphaeifolia.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Hernandiaceae; Hernandia.
OX NCBI_TaxID=121082;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Rohrer J.G.;
RT "Toward a phylogenetic classification of the Lauraceae: Evidence from
RT matK sequences.";
RL Syst. Bot. 25:60-71(2000).
CC -1- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC INTRONS (BY SIMILARITY).
CC -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC MITOCHONDRIAL INTRONS.
CC EMBL: AJ247165; CAC05370.1; -.
CC DR InterPro: IPR000442; Intron_mature2.
CC DR InterPro: IPR002866; MatK_N.
CC DR Pfam: PF01348; Intron_mature2; 1.
CC DR Pfam: PF01824; MatK_N; 1.
CC KW mRNA processing; Chloroplast.
CC FT NON_TER 1
CC FT 360
CC SQ SEQUENCE 360 AA; 42433 MW; AEA86AB324B6A29D CRC64;

Query Match 57.7%; Score 41; DB 8; Length 360;
Best Local Similarity 88.9%; Pred. No. 33;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 7 FLFLSNSFL 15
Db 58 FLFLSNSFLV 66

RESULT 8
O9W707 PRELIMINARY; PRT; 372 AA.
AC O9W707;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE XFD-13; protein.
GN XFD-13.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Koester M., Dillinger K., Knoechel W.;
RT "Genomic structure and embryonic expression of the Xenopus winged
RT helix factors XFD-13/13'.";
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ242679; CAB4731.1; -.
DR HSP; O63245; 2HFH.
DR InterPro: IPR001766; TF_Fork_head.
DR Pfam: PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR ProDom: PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS50039; FORK_HEAD_3; 1.
FT NON_TER 1
FT 372
FT SQ SEQUENCE 372 AA; 40290 MW; 388089A07F1C25CC CRC64;

Query Match 57.7%; Score 41; DB 13; Length 372;
Best Local Similarity 64.3%; Pred. No. 34;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IDLASEFLSNSF 14
Db 11 IDLASEFLSNSF 14

```

```

Db      130 IDPASEFMFEBSGF 143

RESULT 9
ID      Q9W706      PRELIMINARY;      PRT;      373 AA.
AC      Q9W706;
DT      01-NOV-1999 (TrEMBLrel. 12, Created)
DT      01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE      01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
GN      XFD-13 Protein.
OS      Xenopus laevis (African clawed frog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC      Xenopodinae; Xenopus.
OX      NCBI_TaxID=8335;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Koester M., Dillinger K., Knoechel W.;
RT      "Genomic structure and embryonic expression of Xenopus winged helix
RT      factors XFD-13/13.";
RL      Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AJ242680; CAB44732.1; -.
DR      HSSP; O63245; 2HFH.
DR      InterPro; IPR001766; TF_Fork_head.
DR      Pfam; PF00250; Fork_head; 1.
DR      PRINTS; PR00053; FORKHEAD.
DR      ProDom; PD000425; TF_Fork_head; 1.
DR      SMART; SM00339; FH; 1.
DR      PROSITE; PS00657; FORK_HEAD_1; 1.
DR      PROSITE; PS00658; FORK_HEAD_2; 1.
DR      PROSITE; PS50039; FORK_HEAD_3; 1.
SQ      SEQUENCE 373 AA; 40357 MW; 040EEB7BAD63A896 CRC64;

Query Match
Best Local Similarity 57.7%; Score 41; DB 13; Length 373;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 IDLASEFLPLNSNF 14
DB      130 IDPASEFMFEBSGF 143

RESULT 10
ID      O54743      PRELIMINARY;      PRT;      446 AA.
AC      O54743;
DT      01-JUN-1998 (TrEMBLrel. 06, Created)
DT      01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT      01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE      LUN (LUN protein).
GN      FOXF2 OR FOX20 OR LUN.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Miura N., Kakinuma H., Sato M., Alba N., Sugiyama T.;
RL      Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN      [2]
SQ      SEQUENCE OF 11-446 FROM N.A.
RA      Miura N., Kakinuma K., Sato M., Alba N., Sugiyama T.;
RL      Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR      EMBL; Y12293; CAA72972.1; -.
DR      EMBL; Y12294; CAA72972.1; JOINED.
DR      EMBL; Y11148; CAA72035.1; -.
DR      HSSP; O63245; 2HFH.
DR      TRANSFAC; T02514; -.
DR      MGD; MGI:1347479; Foxf2.
DR      InterPro; IPR001766; TF_Fork_head.
DR      Pfam; PF00250; Fork_head; 1.

```

```

DR      PRINTS; PR00053; FORKHEAD.
DR      ProDom; PD000425; TF_Fork_head; 1.
DR      SMART; SM00339; FH; 1.
DR      PROSITE; PS00657; FORK_HEAD_1; 1.
DR      PROSITE; PS00658; FORK_HEAD_2; 1.
DR      PROSITE; PS50039; FORK_HEAD_3; 1.
SQ      SEQUENCE 446 AA; 46374 MW; 472303F238C016A7 CRC64;

Query Match
Best Local Similarity 57.7%; Score 41; DB 11; Length 446;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 IDLASEFLPLNSNF 14
DB      176 IDPASEFMFEBSGF 189

RESULT 11
ID      O63861      PRELIMINARY;      PRT;      164 AA.
AC      O63861;
DT      01-AUG-1998 (TrEMBLrel. 07, Created)
DT      01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT      01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE      NADH dehydrogenase subunit 4 (Fragment).
GN      ND4.
OS      Heterorhabditis sp.
OC      Mitochondrion.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Heterorhabditidae; Heterorhabditis.
OX      NCBI_TaxID=52063;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=99389049; Pubmed=10461953;
RX      Liu J., Berry R.E., Blouin M.S.;
RT      "Molecular differentiation and phylogeny of entomopathogenic nematodes
RT      (Rhabditida: heterorhabditidae) based on ND4 gene sequences of
RT      Mitochondrial DNA.";
RL      J. Parasitol. 85:709-715 (1999).
DR      EMBL; AF066886; AAC19107.1; -.
DR      InterPro; IPR001750; Oxidored_q1.
DR      Pfam; PF00361; oxidored_q1; 1.
KW      NAD; Oxidoreductase; Ubiquinone; Mitochondrion.
FT      NON_TER      1      164
FT      NON_TER      1      164
SQ      SEQUENCE 164 AA; 18675 MW; 2F4E89A4EB2C00EE CRC64;

Query Match
Best Local Similarity 56.3%; Score 40; DB 8; Length 164;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1 IDLASEFLPLNSNF 15
DB      130 LSFISEFMISNGFL 144

RESULT 12
ID      O63862      PRELIMINARY;      PRT;      164 AA.
AC      O63862;
DT      01-AUG-1998 (TrEMBLrel. 07, Created)
DT      01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT      01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE      NADH dehydrogenase subunit 4 (Fragment).
GN      ND4.
OS      Heterorhabditis sp.
OC      Mitochondrion.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Heterorhabditidae; Heterorhabditis.
OX      NCBI_TaxID=52063;
RN      [1]
RP      SEQUENCE FROM N.A.

```



```

RC STRAIN=OH25;
RX MEDLINE=99389049; PubMed=10461953;
RA Liu J., Berry R.E., Blouin M.S.;
RT "Molecular differentiation and phylogeny of entomopathogenic nematodes
RT (rhadditidae: heterorhabditidae) based on ND4 gene sequences of
RT Mitochondrial DNA.";
RL J. Parasitol. 85:709-715(1999).
DR EMBL; AF066887; AAC19108.1; -.
DR InterPro; IPR001750; Oxidored_q1.
KW Pfam; PF00361; oxidored_q1; 1.
KW NAD; Oxidoreductase; Ubiquinone; Mitochondrion.
FT NON TER 1 164
FT SEQUENCE 164 AA; 18597 MW; 62E28443F7607625 CRC64;

Query Match 56.3%; Score 40; DB 8; Length 164;
Best Local Similarity 46.7%; Pred. No. 23;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IDLASEFLPLNSFL 15
Db 130 LSFISFMTISNGFL 144

RESULT 13
ID 063860 PRELIMINARY; PRT; 164 AA.
AC 063860;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE NADH dehydrogenase subunit 4 (Fragment).
GN ND4.
OS Heterorhabditis megidis.
OC Mitochondrion.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhadditida; Rhadditoidae;
OC Heterorhabditidae; Heterorhabditis.
OX NCBI_TaxID=52065;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=H01;
RX MEDLINE=99389049; PubMed=10461953;
RA Liu J., Berry R.E., Blouin M.S.;
RT "Molecular differentiation and phylogeny of entomopathogenic nematodes
RT (rhadditidae: heterorhabditidae) based on ND4 gene sequences of
RT Mitochondrial DNA.";
RL J. Parasitol. 85:709-715(1999).
DR EMBL; AF066885; AAC19106.1; -.
DR InterPro; IPR001750; Oxidored_q1.
KW Pfam; PF00361; oxidored_q1; 1.
KW NAD; Oxidoreductase; Ubiquinone; Mitochondrion.
FT NON TER 1 164
FT SEQUENCE 164 AA; 18675 MW; 2F4E89A4EB2C00EB CRC64;

Query Match 56.3%; Score 40; DB 8; Length 164;
Best Local Similarity 46.7%; Pred. No. 23;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IDLASEFLPLNSFL 15
Db 130 LSFISFMTISNGFL 144

RESULT 14
ID 098650 PRELIMINARY; PRT; 186 AA.
AC 098650;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE Rgg (Fragment).
GN RGG.

```

```

OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=NZ131;
RX MEDLINE=99184993; PubMed=10085009;
RA Chaussee M.S., Ajdic D., Ferretti J.U.;
RT "The rgg gene of Streptococcus pyogenes NZ131 positively influences
RT extracellular SPE B production.";
RL Infect. Immun. 67:1715-1722(1999).
DR EMBL; AF091252; AAD23950.1; -.
FT NON TER 186 186
FT SEQUENCE 186 AA; 22208 MW; B3953946184017ED CRC64;

Query Match 56.3%; Score 40; DB 2; Length 186;
Best Local Similarity 77.8%; Pred. No. 26;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 6 EFLPLNSNF 14
Db 59 EFLPLNSNF 67

RESULT 15
ID 09HFU7 PRELIMINARY; PRT; 245 AA.
AC 09HFU7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Ornithine decarboxylase antizyme.
GN ANTIZYME.
OS Pneumocystis carinii.
OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OX NCBI_TaxID=4754;
RN (1)
RP SEQUENCE FROM N.A.
RA Ivanov I.P., Gesteland R.F., Atkins J.F.;
RT "Antizyme expression: a subversion of triplet decoding, which is
RT remarkably conserved by evolution, is a sensor for an autoregulatory
RT circuit.";
RL Nucleic Acids Res. 28:0-0(2000).
DR EMBL; AF291574; AAG16234.1; -.
SQ SEQUENCE 245 AA; 27677 MW; 2BD98BA35CD5EFBD CRC64;

Query Match 56.3%; Score 40; DB 3; Length 245;
Best Local Similarity 53.3%; Pred. No. 34;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IDLASEFLPLNSFL 15
Db 190 IDPLASFLPCNKLFI 204

```

Search completed: November 26, 2003, 12:34:39
Job time : 34.7108 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:26:20 ; Search time 27.1084 Seconds
(without alignments)
102.059 Million cell updates/sec

Title: US-09-230-111C-16
Perfect score: 73
Sequence: 1 DSEMYNFRSQSLASV 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 673684 seqs, 184443283 residues

Total number of hits satisfying chosen parameters: 673684

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubppaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubppaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubppaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	100.0	15	8	US-08-681-219-18
2	73	100.0	15	12	US-10-092-138-16
3	73	100.0	15	12	US-09-230-111C-16
4	41	56.2	15	8	US-08-681-219-7
5	41	56.2	15	8	US-08-681-219-25
6	41	56.2	15	12	US-10-092-138-5
7	41	56.2	15	12	US-09-230-111C-23
8	41	56.2	15	12	US-09-230-111C-5
9	41	56.2	15	12	US-09-230-111C-23
10	41	56.2	19	12	US-10-341-967-119
11	41	56.2	281	10	US-09-756-854-3
12	41	56.2	281	14	US-10-041-574-3
13	41	56.2	331	12	US-10-280-047-3
14	41	56.2	335	9	US-09-826-212-7
15	41	56.2	335	9	US-09-802-669-2

16	41	56.2	335	9	US-09-333-966-6	Sequence 6, Appl 1
17	41 <th>56.2</th> <th>335</th> <th>9</th> <th>US-09-949-713-20</th> <th>Sequence 20, Appl 1</th>	56.2	335	9	US-09-949-713-20	Sequence 20, Appl 1
18	41 <th>56.2</th> <th>335</th> <th>9</th> <th>US-09-874-138-4</th> <th>Sequence 2, Appl 1</th>	56.2	335	9	US-09-874-138-4	Sequence 2, Appl 1
19	41 <th>56.2</th> <th>335</th> <th>10</th> <th>US-09-884-987-2</th> <th>Sequence 4, Appl 1</th>	56.2	335	10	US-09-884-987-2	Sequence 4, Appl 1
20	41 <th>56.2</th> <th>335</th> <th>10</th> <th>US-09-935-727-9</th> <th>Sequence 9, Appl 1</th>	56.2	335	10	US-09-935-727-9	Sequence 9, Appl 1
21	41 <th>56.2</th> <th>335</th> <th>11</th> <th>US-09-314-889-6</th> <th>Sequence 6, Appl 1</th>	56.2	335	11	US-09-314-889-6	Sequence 6, Appl 1
22	41 <th>56.2</th> <th>335</th> <th>12</th> <th>US-10-189-189-6</th> <th>Sequence 4, Appl 1</th>	56.2	335	12	US-10-189-189-6	Sequence 4, Appl 1
23	41 <th>56.2</th> <th>335</th> <th>13</th> <th>US-10-005-842-4</th> <th>Sequence 3, Appl 1</th>	56.2	335	13	US-10-005-842-4	Sequence 3, Appl 1
24	41 <th>56.2</th> <th>335</th> <th>15</th> <th>US-10-175-902-3</th> <th>Sequence 3, Appl 1</th>	56.2	335	15	US-10-175-902-3	Sequence 3, Appl 1
25	41 <th>56.2</th> <th>335</th> <th>15</th> <th>US-10-186-643-7</th> <th>Sequence 7, Appl 1</th>	56.2	335	15	US-10-186-643-7	Sequence 7, Appl 1
26	41 <th>56.2</th> <th>335</th> <th>15</th> <th>US-10-226-296-3</th> <th>Sequence 3, Appl 1</th>	56.2	335	15	US-10-226-296-3	Sequence 3, Appl 1
27	41 <th>56.2</th> <th>669</th> <th>15</th> <th>US-10-226-318-3</th> <th>Sequence 25, Appl 1</th>	56.2	669	15	US-10-226-318-3	Sequence 25, Appl 1
28	41 <th>56.2</th> <th>681</th> <th>10</th> <th>US-09-969-384-25</th> <th>Sequence 140, Appl 1</th>	56.2	681	10	US-09-969-384-25	Sequence 140, Appl 1
29	41 <th>56.2</th> <th>681</th> <th>12</th> <th>US-10-158-825-140</th> <th>Sequence 140, Appl 1</th>	56.2	681	12	US-10-158-825-140	Sequence 140, Appl 1
30	41 <th>56.2</th> <th>681</th> <th>15</th> <th>US-10-158-847-140</th> <th>Sequence 13, Appl 1</th>	56.2	681	15	US-10-158-847-140	Sequence 13, Appl 1
31	41 <th>56.2</th> <th>711</th> <th>10</th> <th>US-09-969-384-13</th> <th>Sequence 138, Appl 1</th>	56.2	711	10	US-09-969-384-13	Sequence 138, Appl 1
32	41 <th>56.2</th> <th>711</th> <th>12</th> <th>US-10-158-825-138</th> <th>Sequence 138, Appl 1</th>	56.2	711	12	US-10-158-825-138	Sequence 138, Appl 1
33	41 <th>56.2</th> <th>711</th> <th>15</th> <th>US-10-158-847-138</th> <th>Sequence 2, Appl 1</th>	56.2	711	15	US-10-158-847-138	Sequence 2, Appl 1
34	41 <th>56.2</th> <th>805</th> <th>10</th> <th>US-09-978-385-2</th> <th>Sequence 142, Appl 1</th>	56.2	805	10	US-09-978-385-2	Sequence 142, Appl 1
35	41 <th>56.2</th> <th>805</th> <th>12</th> <th>US-10-158-825-142</th> <th>Sequence 86, Appl 1</th>	56.2	805	12	US-10-158-825-142	Sequence 86, Appl 1
36	41 <th>56.2</th> <th>805</th> <th>14</th> <th>US-10-114-893-86</th> <th>Sequence 142, Appl 1</th>	56.2	805	14	US-10-114-893-86	Sequence 142, Appl 1
37	41 <th>56.2</th> <th>805</th> <th>15</th> <th>US-10-158-847-142</th> <th>Sequence 570, Appl 1</th>	56.2	805	15	US-10-158-847-142	Sequence 570, Appl 1
38	41 <th>56.2</th> <th>805</th> <th>15</th> <th>US-10-005-956-570</th> <th>Sequence 843, Appl 1</th>	56.2	805	15	US-10-005-956-570	Sequence 843, Appl 1
39	41 <th>56.2</th> <th>805</th> <th>15</th> <th>US-10-005-956-843</th> <th>Sequence 731, Appl 1</th>	56.2	805	15	US-10-005-956-843	Sequence 731, Appl 1
40	38 <th>52.1</th> <th>524</th> <th>9</th> <th>US-09-764-853-731</th> <th>Sequence 202, Appl 1</th>	52.1	524	9	US-09-764-853-731	Sequence 202, Appl 1
41	38 <th>52.1</th> <th>524</th> <th>15</th> <th>US-10-091-438-202</th> <th>Sequence 494, Appl 1</th>	52.1	524	15	US-10-091-438-202	Sequence 494, Appl 1
42	38 <th>52.1</th> <th>654</th> <th>9</th> <th>US-09-764-853-494</th> <th>Sequence 36, Appl 1</th>	52.1	654	9	US-09-764-853-494	Sequence 36, Appl 1
43	38 <th>52.1</th> <th>944</th> <th>12</th> <th>US-10-335-592-3</th> <th>Sequence 24, Appl 1</th>	52.1	944	12	US-10-335-592-3	Sequence 24, Appl 1
44	38 <th>52.1</th> <th>1694</th> <th>12</th> <th>US-10-203-708-36</th> <th></th>	52.1	1694	12	US-10-203-708-36	
45	38 <th>52.1</th> <th>1700</th> <th>12</th> <th>US-09-863-776-24</th> <th></th>	52.1	1700	12	US-09-863-776-24	

ALIGNMENTS

RESULT 1
US-08-681-219-18
; Sequence 18, Application US/08681219
; Publication No. US20020058607A1
; GENERAL INFORMATION:
; APPLICANT: Takaaki Sato and Junn Yanagisawa
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GIGP
; NUMBER OF INVENTIONS: (PDE/DHR) DOMAIN AND USBS THEREOF
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/681,219
; FILING DATE: 22-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/48962/JPW/JMK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-681-219-18

Query Match 100.0%; Score 73; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSEMYNFRSOLASV 15
DB 1 DSEMYNFRSOLASV 15

RESULT 2
US-10-092-138-16
Sequence 16, Application US/10092138
Publication No. US20030170723A1
GENERAL INFORMATION:

APPLICANT: Sato, Taka-Aki
TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
FILE REFERENCE: 65823/JPW/PT
CURRENT APPLICATION NUMBER: US/10/092,138
CURRENT FILING DATE: 2002-09-06
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 16
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial
US-10-092-138-16

Query Match 100.0%; Score 73; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSEMYNFRSOLASV 15
DB 1 DSEMYNFRSOLASV 15

RESULT 3
US-09-230-111C-16
Sequence 16, Application US/09230111C
Publication No. US20030203414A1
GENERAL INFORMATION:

APPLICANT: Yanoigisawa, Junn
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN
TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF (PDZ/DHR)
FILE REFERENCE: 48962-A-PCT-US
CURRENT APPLICATION NUMBER: US/09/230,111C
CURRENT FILING DATE: 1999-05-17
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 16
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial
US-09-230-111C-16

Query Match 100.0%; Score 73; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSEMYNFRSOLASV 15

DB 1 DSEMYNFRSOLASV 15

RESULT 4
US-08-681-219-7
Sequence 7, Application US/08681219
Publication No. US20020058607A1
GENERAL INFORMATION:

APPLICANT: Takaaki Sato and Junn Yanoigisawa
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF
TITLE OF INVENTION: (PDZ/DHR) DOMAIN AND USES THEREOF
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/681,219
FILING DATE: 22-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/48962/JPW/JKM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-681-219-7

Query Match 56.2%; Score 41; DB 8; Length 15;
Best Local Similarity 53.3%; Pred. No. 0.4;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSOLASV 15
DB 1 DSEMYNFRSOLASV 15

RESULT 5
US-08-681-219-25
Sequence 25, Application US/08681219
Publication No. US20020058607A1
GENERAL INFORMATION:
APPLICANT: Takaaki Sato and Junn Yanoigisawa
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF
TITLE OF INVENTION: (PDZ/DHR) DOMAIN AND USES THEREOF
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:

```
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/681,219
/ FILING DATE: 22-JUL-1996
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: White, John P
/ REGISTRATION NUMBER: 28,678
/ REFERENCE/DOCKET NUMBER: 0575/48962/JPM/JKM
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 278-0400
/ TELEFAX: (212) 391-0525
/ INFORMATION FOR SEQ ID NO: 25:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 15 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-08-681-219-25

Query Match          56.2%; Score 41; DB 8; Length 15;
Best Local Similarity 53.3%; Pred. No. 0.4;
Matches      8; Conservative 4; Mismatches      3; Indels 0; Gaps 0;

QY      1 DSEMYNFRSQSLASV 15
DB      1 DSENSNFRNEIQSLV 15

RESULT 6
US-10-092-138-5
/ Sequence 5, Application US/10092138
/ Publication No. US20030170723A1
/ GENERAL INFORMATION:
/ APPLICANT: Sato, Taka-Aki
/ TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
/ TITLE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION
/ FILE REFERENCE: 65823/JPM/PT
/ CURRENT APPLICATION NUMBER: US/10/092,138
/ CURRENT FILING DATE: 2002-09-06
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 5
/ LENGTH: 15
/ TYPE: PRT
/ ORGANISM: human
US-10-092-138-5

Query Match          56.2%; Score 41; DB 12; Length 15;
Best Local Similarity 53.3%; Pred. No. 0.4;
Matches      8; Conservative 4; Mismatches      3; Indels 0; Gaps 0;

QY      1 DSEMYNFRSQSLASV 15
DB      1 DSENSNFRNEIQSLV 15

RESULT 7
US-10-092-138-23
/ Sequence 23, Application US/10092138
/ Publication No. US20030170723A1
/ GENERAL INFORMATION:
/ APPLICANT: Sato, Taka-Aki
/ TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
/ TITLE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION
/ FILE REFERENCE: 65823/JPM/PT
/ CURRENT APPLICATION NUMBER: US/10/092,138
/ CURRENT FILING DATE: 2002-09-06
/ NUMBER OF SEQ ID NOS: 33
```

```
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 23
/ LENGTH: 15
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial
US-10-092-138-23

Query Match          56.2%; Score 41; DB 12; Length 15;
Best Local Similarity 53.3%; Pred. No. 0.4;
Matches      8; Conservative 4; Mismatches      3; Indels 0; Gaps 0;

QY      1 DSEMYNFRSQSLASV 15
DB      1 DSENSNFRNEIQSLV 15

RESULT 8
US-09-230-111C-5
/ Sequence 5, Application US/09230111C
/ Publication No. US20030203414A1
/ GENERAL INFORMATION:
/ APPLICANT: Sato, Taka-Aki
/ APPLICANT: Yanagisawa, Junn
/ TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN
/ TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF (PDZ/DHR)
/ FILE REFERENCE: 48962-A-PCT-US
/ CURRENT APPLICATION NUMBER: US/09/230,111C
/ CURRENT FILING DATE: 1999-05-17
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 5
/ LENGTH: 15
/ TYPE: PRT
/ ORGANISM: human
US-09-230-111C-5

Query Match          56.2%; Score 41; DB 12; Length 15;
Best Local Similarity 53.3%; Pred. No. 0.4;
Matches      8; Conservative 4; Mismatches      3; Indels 0; Gaps 0;

QY      1 DSEMYNFRSQSLASV 15
DB      1 DSENSNFRNEIQSLV 15

RESULT 9
US-09-230-111C-23
/ Sequence 23, Application US/09230111C
/ Publication No. US20030203414A1
/ GENERAL INFORMATION:
/ APPLICANT: Sato, Taka-Aki
/ APPLICANT: Yanagisawa, Junn
/ TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN
/ TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF (PDZ/DHR)
/ FILE REFERENCE: 48962-A-PCT-US
/ CURRENT APPLICATION NUMBER: US/09/230,111C
/ CURRENT FILING DATE: 1999-05-17
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 23
/ LENGTH: 15
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial
US-09-230-111C-23
```

Query Match 56.2%; Score 41; DB 12; Length 15;
Best Local Similarity 53.3%; Pred. No. 0.4;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSQSLASV 15
||| |||:::|
1 DSENSNFRNEIOSLV 15

Db 1 DSENSNFRNEIOSLV 15

RESULT 10
US-10-341-967-119
; Sequence 119, Application US/10341967
; Publication No. US20030206891A1
; GENERAL INFORMATION:
; APPLICANT: ARIAD Gene Therapeutics, Inc.
; TITLE OF INVENTION: Rapamycin Based Regulation of Biological Events
; FILE REFERENCE: 345B PCT
; CURRENT APPLICATION NUMBER: US/10/341,967
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US/09/481,620A
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 119
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Spel mutation in FAS
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(119)
US-10-341-967-119

Query Match 56.2%; Score 41; DB 12; Length 19;
Best Local Similarity 53.3%; Pred. No. 0.52;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSQSLASV 15
||| |||:::|
3 DSENSNFRNEIOSLV 17

Db 3 DSENSNFRNEIOSLV 17

RESULT 11
US-09-756-854-3
; Sequence 3, Application US/09756854
; Patent No. US20020164684A1
; GENERAL INFORMATION:
; APPLICANT: NI, Jian
; Yu, Guo-Liang
; Fan, Ping
; Gentz, Reiner
; TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor TR9
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: US
; ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/756,854
FILING DATE: 10-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/095,094
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Hoover, Kenley K.
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PF375
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 281 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-756-854-3

Query Match 56.2%; Score 41; DB 10; Length 281;
Best Local Similarity 53.3%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSQSLASV 15
||| |||:::|
267 DSENSNFRNEIOSLV 281

Db 267 DSENSNFRNEIOSLV 281

RESULT 12
US-10-041-574-3
; Sequence 3, Application US/10041574
; Publication No. US20020168359A1
; GENERAL INFORMATION:
; APPLICANT: NI, Jian
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Fan, Ping
; APPLICANT: Gentz, Reiner L.
; TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor TR9
; FILE REFERENCE: PF375PI
; CURRENT APPLICATION NUMBER: US/10/041,574
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: 09/527,236
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/052,991
; PRIOR FILING DATE: 1997-06-11
; PRIOR APPLICATION NUMBER: 09/095,094
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/126,019
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: 60/134,220
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-041-574-3

Query Match 56.2%; Score 41; DB 14; Length 281;
Best Local Similarity 53.3%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSQSLASV 15
||| |||:::|
267 DSENSNFRNEIOSLV 281

Db 267 DSENSNFRNEIOSLV 281

RESULT 13
US-10-280-047-3
; Sequence 3, Application US/10280047
; Publication No. US2003018083A1
; GENERAL INFORMATION:
; APPLICANT: NI, Jian
; APPLICANT: Rosen, Craig A

```

; TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor 10
; FILE REFERENCE: P3799PDI
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 09/580,212
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/086,483
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/069,112
; PRIOR FILING DATE: 1997-12-09
; PRIOR APPLICATION NUMBER: 60/050,936
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/144,023
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/142,563
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/136,786
; PRIOR FILING DATE: 1999-05-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 331
; TYPE: PRT
; ORGANISM: human
US-10-280-047-3

Query Match      56.2% Score 41; DB 12; Length 331;
Best Local Similarity 53.3% Pred. No. 15;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy      1 DSEMYNFRSQSLASV 15
      ||| |||::: |||
Db      317 DSENSNFRNFIQSLV 331

RESULT 14
US-09-826-212-7
; Sequence 7, Application US/09826212
; Patent No. US20010021516A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Ying-Fei
; APPLICANT: Gentz, Reiner
; APPLICANT: Ruben, Steven
; APPLICANT: Ni, Jian
; TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
; FILE REFERENCE: 1488,1280006
; CURRENT APPLICATION NUMBER: US/09/826,212
; CURRENT FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-212-7

Query Match      56.2% Score 41; DB 9; Length 335;
Best Local Similarity 53.3% Pred. No. 15;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy      1 DSEMYNFRSQSLASV 15
      ||| |||::: |||
Db      321 DSENSNFRNFIQSLV 335

RESULT 15
US-09-802-669-2
; Sequence 2, Application US/09802669
; Patent No. US20020004490A1
; GENERAL INFORMATION:
; APPLICANT: Dean, Nicholas M.
; APPLICANT: Marcussen, Eric G.
; APPLICANT: Wyatt, Jacqueline
```

```

; APPLICANT: Zhang, Hong
; TITLE OF INVENTION: Antisense Compound Modulation of Fas Mediated Signaling
; FILE REFERENCE: ISFH-545
; CURRENT APPLICATION NUMBER: US/09/802,669
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 09/665,615
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 09/290,640
; PRIOR FILING DATE: 1999-04-12
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-802-669-2

Query Match      56.2% Score 41; DB 9; Length 335;
Best Local Similarity 53.3% Pred. No. 15;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy      1 DSEMYNFRSQSLASV 15
      ||| |||::: |||
Db      321 DSENSNFRNFIQSLV 335
```

Search completed: November 26, 2003, 12:38:40
Job time : 27.1084 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:21:29 ; Search time 41.5663 Seconds
(without alignment)

57.280 Million cell updates/sec

Title: US-09-230-111c-16

Perfect score: 73
Sequence: 1 DSEMYNFRSQLASV 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19jun03:*

- 1: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	100.0	15	19	AAW50173
2	41	56.2	15	18	AAW25604
3	41	56.2	20	22	AAW55801
4	41	56.2	20	22	AAW57663
5	41	56.2	20	22	AAW58073
6	41	56.2	20	23	ABJ05193
7	41	56.2	20	23	ABP63211
8	41	56.2	121	23	ABB81752
9	41	56.2	281	21	AAW26982

Result No.	Score	Query Match	Length	DB ID	Description
10	41	56.2	314	15	AAW76238
11	41	56.2	314	17	AAW96682
12	41	56.2	314	20	AAW98070
13	41	56.2	331	22	AAW50893
14	41	56.2	335	13	AAW28084
15	41	56.2	335	16	AAW78606
16	41	56.2	335	17	AAW9681
17	41	56.2	335	17	AAW92528
18	41	56.2	335	17	AAW92528
19	41	56.2	335	19	AAW49104
20	41	56.2	335	21	AAW19341
21	41	56.2	335	21	AAW36267
22	41	56.2	335	21	AAW31335
23	41	56.2	335	22	AAW50517
24	41	56.2	335	24	AAW07407
25	41	56.2	431	22	AAW97650
26	41	56.2	431	24	ABU04641
27	41	56.2	436	21	AAW91026
28	41	56.2	436	22	AAW97651
29	41	56.2	436	24	ABU04642
30	41	56.2	669	19	AAW64484
31	41	56.2	711	22	AAW09092
32	41	56.2	805	21	AAW84562
33	41	56.2	805	21	AAW67310
34	41	56.2	805	22	AAW72667
35	41	56.2	805	22	AAW48095
36	41	56.2	805	23	ABG77011
37	41	56.2	805	23	ABG77023
38	41	56.2	805	23	AAW99701
39	41	56.2	805	23	AAE20353
40	41	56.2	805	24	ABU07731
41	41	56.2	920	22	AAW97653
42	41	56.2	927	22	AAW97653
43	41	56.2	1066	17	AAW91300
44	41	56.2	1066	19	AAW33632
45	40	54.8	193	22	AAW66326

ALIGNMENTS

RESULT 1
ID AAW50173 standard; peptide; 15 AA.
AC AAW50173;
XX
DT 16-JUL-1998 (first entry)
XX
DE Signal-transducing protein carboxy-terminal peptide.
XX
KW Inhibition; specific binding; signal-transducing protein;
KW cytoplasmic protein; proliferation; cancer cell; apoptosis;
KW virally infected cell.
XX
OS Synthetic.
XX
PN WO9805347-A1.
XX
PD 12-FEB-1998.
XX
PF 18-JUL-1997; 97WO-US12677.
XX
PR 22-JUL-1996; 96US-0681219.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Sato T, Yanagisawa J;
XX
DR WPI, 1998-145347/13.
XX
PT Inhibition of signal transduction - by inhibiting binding between a
PT signal-transducing protein and a cytoplasmic protein, for treating

```

PT e.g. cancer or viral infection
XX
XX Claim 15; Page 64; 108pp; English.
XX
XX A novel composition is capable of inhibiting specific binding
CC between a signal-transducing protein (STP) having the
CC carboxy-terminal sequence (Ser/Thr)-Xaa-(Val/Ile/Leu), where Xaa =
CC any amino acid (e.g. the present peptide), and a cytoplasmic
CC protein (CP) containing the sequence AAW50162 or AAW50163.
CC The composition can be used to inhibit the proliferation of cancer
CC or virally infected cells, or induce apoptosis in cancer or virally
CC infected cells.
XX
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 19; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSEMYNFRSOLASV 15
DB 1 DSEMYNFRSOLASV 15

RESULT 2
AAW25604
ID AAW25604 standard; peptide; 15 AA.
XX
XX AAW25604;
AC
XX
XX 03-NOV-1997 (first entry)
DT
XX
XX Fas-antigen.
DE
XX
XX Antigen; immunological assay; Fas; anti-Fas antibody.
XX
XX Synthetic.
XX
XX JP09166593-A.
XX
XX 24-JUN-1997.
PD
XX
XX 14-DEC-1995; 95JP-0325680.
PF
XX
XX 14-DEC-1995; 95JP-0325680.
PR
XX
XX (IGAK-) IGAKU SEIBUTSUGAKU KENKYUSHO KK.
PA
XX
XX WPI; 1997-382345/35.
DR
XX
XX Kit for immunological assay of Fas antigen - comprises anti-Fas
PT antibody specifically bound to area of cell of Fas antigen
XX
XX Claim 2; Page 7; 9pp; Japanese.
XX
XX The sequences given in AAW25604-05 are antigens which are used in the
CC kits of the invention. The kits allow for immunological assay of Fas
CC antigen. They comprise anti-Fas antibody specifically bound to an area of
CC a cell showing Fas antigen, which is immobilised on an insoluble
CC supporting material. The labelled anti-Fas antibody specifically binds
CC to "out-of-area" of Fas antigen and standard Fas antigen extracted from
CC cells.
XX
XX
SQ Sequence 15 AA;

Query Match 56.2%; Score 41; DB 18; Length 15;
Best Local Similarity 53.3%; Pred. No. 0.55;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSOLASV 15
DB 1 DSEMYNFRSOLASV 15

```

```

RESULT 3
AAB55801
ID AAB55801 standard; Peptide; 20 AA.
XX
XX AAB55801;
AC
XX
XX 07-MAR-2001 (first entry)
DT
XX
XX PL peptide #12.
DE
XX
XX Hematopoietic cell; PDZ; PL; autoimmune disease; inflammation;
KW allergy; asthma; multiple sclerosis; cancer; infection.
XX
XX Synthetic.
XX
XX WO200069896-A2.
XX
XX 23-NOV-2000.
PD
XX
XX 12-MAY-2000; 2000MO-US13161.
PF
XX
XX 14-MAY-1999; 99US-0134114.
PR
XX
XX 14-MAY-1999; 99US-0134117.
PR
XX
XX 14-MAY-1999; 99US-0134118.
PR
XX
XX 21-OCT-1999; 99US-0160860.
PR
XX
XX 29-OCT-1999; 99US-0162498.
PR
XX
XX 13-DEC-1999; 99US-0170453.
PR
XX
XX 14-JAN-2000; 2000US-0176195.
PR
XX
XX 14-FEB-2000; 2000US-0182296.
PR
XX
XX 11-APR-2000; 2000US-0196460.
PR
XX
XX 11-APR-2000; 2000US-0196527.
XX
XX (ARBO-) ARBOR VITA CORP.
XX
XX Lu PS;
XX
XX WPI; 2001-080245/09.
DR
XX
XX Modulating a biological function of an endothelial cell or
PT hematopoietic cell, useful for treating autoimmune diseases and
PT infectious diseases, by administering an antagonist that inhibits
PT binding between a PDZ protein and a PL protein -
XX
XX Disclosure; Page 53; 141pp; English.
XX
XX The present invention relates to a new method for modulating a
CC biological function of an endothelial cell or hematopoietic cell. The
CC method involves introducing into a cell, an antagonist that inhibits
CC binding between a PDZ protein and a PL protein. The inhibitor is used
CC to treat a disease mediated by hematopoietic cells, e.g. autoimmune
CC disease. It may also be used to prevent transplantation rejection of
CC a solid organ transplant. The method may also be used in the treatment
CC of inflammation, allergy, inflammatory bowel diseases, ulcerative
CC colitis, ileitis, psoriasis, asthma, atopic dermatitis, autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis,
CC insulin-dependent diabetes, Hashimoto thyroiditis, osteoarthritis,
CC graft rejection, transplantation rejection), atherosclerosis, cancers,
CC infectious diseases, ischemia, vasculitis and Crohn's disease.
XX
XX
SQ Sequence 20 AA;

Query Match 56.2%; Score 41; DB 22; Length 20;
Best Local Similarity 53.3%; Pred. No. 0.76;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSOLASV 15
DB 6 DSEMYNFRSOLASV 20

```

RESULT 4
AAB57663

ID AAB57663 standard; Peptide; 20 AA.
 XX AAB57663;
 XX
 XX
 DT 12-MAR-2001 (first entry)
 XX
 DE CD95 PL peptide.
 XX
 KW Endothelial cell; haematopoietic cell; PDZ domain protein;
 KW PL domain protein; leukocyte activation; synapse formation;
 KW transmembrane neurotransmitter receptor; autoimmune disease;
 KW transplantation rejection; inflammation; allergy;
 KW inflammatory bowel disease; ulcerative colitis; ileitis; psoriasis;
 KW asthma; atopic dermatitis; atherosclerosis; cancer; infectious disease;
 KW lechaemia; vasculitis; Crohn's disease.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200069897-A2.
 PN
 XX
 PD 23-NOV-2000.
 XX
 PF 12-MAY-2000; 2000WO-US13166.
 XX
 PR 14-MAY-1999; 99US-0134114.
 XX
 PR 14-MAY-1999; 99US-0134117.
 XX
 PR 14-MAY-1999; 99US-0134118.
 XX
 PR 21-OCT-1999; 99US-0160860.
 XX
 PR 29-OCT-1999; 99US-0162498.
 XX
 PR 13-DEC-1999; 99US-0170453.
 XX
 PR 14-JAN-2000; 2000US-0176195.
 XX
 PR 14-FEB-2000; 2000US-0182296.
 XX
 PR 11-APR-2000; 2000US-0196460.
 XX
 PR 11-APR-2000; 2000US-0196527.
 XX
 PA (ARBO-) ARBOR VITA CORP.
 XX
 XX Lu PS;
 PI
 DR WPI; 2001-025003/03.
 XX
 PT New inhibitors of binding of a PDZ protein and PL protein for
 PT inhibiting T cell-mediated response by hematopoietic cells, or for
 PT treating diseases characterized by inflammatory and humoral immune
 PT responses, e.g. inflammation, cancer -
 XX
 XX
 PS Disclosure; Page 51; 139pp; English.
 XX
 XX
 CC The present invention relates to a method for modulating a biological
 CC function of an endothelial cell or haematopoietic cell, comprises
 CC introducing into a cell an antagonist that inhibits binding between a
 CC PDZ domain protein and a PL domain protein to result in inhibition of
 CC leukocyte activation. The present sequence is a PL peptide. PDZ domains
 CC of proteins are named after three prototypical proteins: PSD95,
 CC Drosophila large disc protein and Zonula Occludin 1 protein. PDZ domain
 CC proteins are involved in synapse formation by organising transmembrane
 CC neurotransmitter receptors through intracellular interactions. The
 CC inhibitors identified by the present invention can be used to treat a
 CC disease mediated by haematopoietic cells, e.g. autoimmune disease,
 CC inflammation, allergy (e.g. drug allergies), inflammatory bowel diseases,
 CC ulcerative colitis, ileitis, psoriasis, respiratory allergic diseases
 CC (e.g. asthma), atopic dermatitis, autoimmune diseases (e.g. rheumatoid
 CC arthritis, multiple sclerosis, insulin-dependent diabetes, Hashimoto
 CC thyroiditis, osteoarthritis), atherosclerosis, cancers, infectious
 CC diseases (e.g. viral infection), lechaemia, vasculitis and Crohn's
 CC disease. The inhibitors can also be used to prevent transplantation
 CC rejection of a solid organ transplant.
 XX
 XX Sequence 20 AA;
 SQ
 Query Match 56.2%; Score 41; DB 22; Length 20;
 Best Local Similarity 53.3%; Pred. No. 0.76;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMTNFSQLASV 15
 ||| |||::|:
 Db 6 DSEMTNFSQLASV 20
 RESULT 5
 ID AAB58073
 XX AAB58073 standard; Peptide; 20 AA.
 AC
 XX
 DT 12-MAR-2001 (first entry)
 XX
 DE CD95 PL peptide.
 XX
 KW Endothelial cell; haematopoietic cell; PDZ domain protein;
 KW PL domain protein; leukocyte activation; synapse formation;
 KW transmembrane neurotransmitter receptor; autoimmune disease;
 KW transplantation rejection; inflammation; allergy;
 KW inflammatory bowel disease; ulcerative colitis; ileitis; psoriasis;
 KW asthma; atopic dermatitis; atherosclerosis; cancer; infectious disease;
 KW lechaemia; vasculitis; Crohn's disease.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200069898-A2.
 PN
 XX
 PD 23-NOV-2000.
 XX
 PF 12-MAY-2000; 2000WO-US13205.
 XX
 PR 14-MAY-1999; 99US-0134114.
 XX
 PR 14-MAY-1999; 99US-0134117.
 XX
 PR 14-MAY-1999; 99US-0134118.
 XX
 PR 21-OCT-1999; 99US-0160860.
 XX
 PR 29-OCT-1999; 99US-0162498.
 XX
 PR 13-DEC-1999; 99US-0170453.
 XX
 PR 14-JAN-2000; 2000US-0176195.
 XX
 PR 14-FEB-2000; 2000US-0182296.
 XX
 PR 11-APR-2000; 2000US-0196460.
 XX
 PR 11-APR-2000; 2000US-0196527.
 XX
 PA (ARBO-) ARBOR VITA CORP.
 XX
 XX Lu PS;
 PI
 DR WPI; 2001-061214/07.
 XX
 PT Modulating a biological function of a hematopoietic cell for treating
 PT an allergic response, or diseases mediated by immune system cells,
 PT comprises introducing into the cell a PDZ-PL interaction enhancer or
 PT inhibitor -
 XX
 XX
 PS Disclosure; Page 54; 143pp; English.
 XX
 XX
 CC The present invention relates to a method for modulating a biological
 CC function of an endothelial cell or haematopoietic cell, comprises
 CC introducing into a cell an antagonist that inhibits binding between a
 CC PDZ domain protein and a PL domain protein to result in inhibition of
 CC leukocyte activation. The present sequence is a PL peptide. PDZ domains
 CC of proteins are named after three prototypical proteins: PSD95,
 CC Drosophila large disc protein and Zonula Occludin 1 protein. PDZ domain
 CC proteins are involved in synapse formation by organising transmembrane
 CC neurotransmitter receptors through intracellular interactions. The
 CC inhibitors identified by the present invention can be used to treat a
 CC disease mediated by haematopoietic cells, e.g. autoimmune disease,
 CC inflammation, allergy (e.g. drug allergies), inflammatory bowel diseases,
 CC ulcerative colitis, ileitis, psoriasis, respiratory allergic diseases
 CC (e.g. asthma), atopic dermatitis, autoimmune diseases (e.g. rheumatoid
 CC arthritis, multiple sclerosis, insulin-dependent diabetes, Hashimoto
 CC thyroiditis, osteoarthritis), atherosclerosis, cancers, infectious
 CC diseases (e.g. viral infection), lechaemia, vasculitis and Crohn's

CC disease. The inhibitors can also be used to prevent transplantation
CC rejection of a solid organ transplant.

XX
SQ Sequence 20 AA;

Query Match 56.2%; Score 41; DB 22; Length 20;
Best Local Similarity 53.3%; Pred. No. 0.76;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYFRSOLASV 15
||| |||::|:
DB 6 DSENSFRNEIOSLV 20

RESULT 6

AB05193
ID AB05193 standard; Peptide; 20 AA.

XX
AC AB05193;

XX
DT 07-NOV-2002 (first entry)

XX
DE C-terminal peptide of a PDZ-binding protein SEQ ID No 188.

XX
KW Immunosuppressive; antiinflammatory; affinity; Kd; binding; PDZ domain;
KW ligand; Ki; inhibitor; K-enhancer; leukocyte; autoimmune disease;
KW inflammatory; humoral immune response; inflammation; C-terminal;
KW PDZ-binding protein.

XX
OS Unidentified.

XX
PN WO200231512-A2.

XX
PD 18-APR-2002.

XX
PF 11-OCT-2001; 2001WO-US32150.

XX
PR 13-OCT-2000; 2000US-0688017.

XX
PA (ARBO-) ARBOR VITA CORP.

XX
PI Rabinowitz JD, Lu PS, Schweizer J;

XX
DR WPI; 2002-416878/44.

XX
PT Assays for determining the affinity of binding between a PDZ domain and
PT a ligand, and determining the Ki of an inhibitor of the binding,
PT comprises using a polypeptide comprising a PDZ domain and a non-PDZ
PT domain -

XX
PS Disclosure; Page 61; 164pp; English.

XX
CC The invention relates to methods and reagents for determining the
CC apparent affinity (Kd) of binding between a PDZ domain and a ligand. The
CC invention also relates to methods and reagents for determining the Ki of
CC an inhibitor of binding between a PDZ domain and a ligand, identifying an
CC agent that enhances binding of a PDZ domain and a ligand, and determining
CC the potency (K-enhancer) of binding between a PDZ domain and a ligand, by
CC determining the ligand bound with an immobilised polypeptide comprising a
CC PDZ domain and a non-PDZ domain on a surface. The modulator (preferably,
CC an inhibitor) of interaction between PDZ and PL is useful for treating a
CC disease characterised by leukocyte activation, e.g., an autoimmune
CC disease that is characterised by inflammatory or humoral immune response,
CC and for reducing inflammation in a subject. This sequence represents a
CC C-terminal peptide of a PDZ-binding protein relating to the invention.

XX
SQ Sequence 20 AA;

Query Match 56.2%; Score 41; DB 23; Length 20;

Best Local Similarity 53.3%; Pred. No. 0.76;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYFRSOLASV 15

DB ||| |||::|:
6 DSENSFRNEIOSLV 20

RESULT 7
ABP63211
ID ABP63211 standard; Peptide; 20 AA.

XX
AC ABP63211;

XX
DT 28-OCT-2002 (first entry)

XX
DE CD95 PL peptide.

XX
KW Molecular interaction; haematopoietic cell; immune response; T cell;
KW PDZ domain; B cell; endothelial cell; PDZ protein; PSD95; PDZ ligand;
KW Drosophila large disc protein; Zonula Occludin 1 protein; PL protein;
KW Immunosuppressive; antiinflammatory; anti-allergic; antithrombotic;
KW antitumor; antiproliferative; dermatological; antiaesthetic; cytostatic;
KW antimicrobial; vasotropic; inflammatory immune response; inflammation;
KW humoral immune response; autoimmune disease; allergy; ulcerative colitis;
KW inflammatory bowel disease; ileitis; enteritis; psoriasis; scleroderma;
KW inflammatory dermatosis; respiratory allergic disease; asthma; cancer;
KW allergic rhinitis; transplantation rejection; atherosclerosis; ischaemia;
KW angiogenesis-dependent disorder; infectious disease.

XX
OS Homo sapiens.

XX
OS Synthetic.

XX
PN WO200242422-A2.

XX
PD 30-MAY-2002.

XX
PF 09-NOV-2001; 2001WO-US44138.

XX
PR 11-NOV-2000; 2000US-0710059.

XX
PR 24-NOV-2000; 2000US-0721915.

XX
PR 24-NOV-2000; 2000US-0722069.

XX
PR 28-NOV-2000; 2000US-0724553.

XX
PA (ARBO-) ARBOR VITA CORP.

XX
PI Lu P, Rabinowitz JD, Schweizer J;

XX
DR WPI; 2002-608221/65.

XX
PT Modulating the biological function of an endothelial cell or
PT hemotopoietic cell e.g., a T-cell or B-cell comprises introducing into
PT the cell, an agent that inhibits binding of a PDZ protein and a PDZ
PT ligand protein in the cell -

XX
PS Disclosure; Page 64; 207pp; English.

XX
CC The present invention describes a method (M1) for modulating a biological
CC function of an endothelial cell or haematopoietic cell. M1 comprises
CC introducing into the cell, an agent that inhibits binding of a PDZ
CC (PS95, Drosophila large disc protein, and Zonula Occludin 1 protein)
CC protein and a PDZ ligand (PL) protein in the cell, and so modulates the
CC biological function. Also described is a method (M2) for determining
CC whether a test compound is an inhibitor of binding between a PDZ protein
CC and a PL protein. M1 is used for modulating a biological function of an
CC endothelial cell or haematopoietic cell e.g., T-cell or B-cell, by an
CC inflammatory or humoral immune response, or an autoimmune disease. An
CC inhibitor (I) is useful for treating a disease characterised by leukocyte
CC activation, where the disease is characterised by an inflammatory or
CC humoral immune response, e.g., an autoimmune disease. The compounds
CC e.g., PL-PDZ interaction inhibitors are useful for treating (ameliorating
CC symptoms of) a variety of diseases and conditions characterised by
CC inflammatory and humoral immune responses e.g., inflammation, allergy,
CC inflammatory bowel diseases, ulcerative colitis, ileitis and enteritis,
CC psoriasis and inflammatory dermatoses, scleroderma, respiratory allergic
CC diseases such as asthma, allergic rhinitis, transplantation rejection
CC (cardiac, kidney, lung, liver, small bowel, pancreas, cadaver,

CC autologous, bone marrow, xenotransplantation), atherosclerosis, cancers,
 CC anglogenesis-dependent disorders, infectious diseases and ischemia.
 CC AB096620 to AB096732 and AB093153 to AB093578 represent sequences used
 CC in the exemplification of the present invention.

XX Sequence 20 AA;

Query Match 56.2%; Score 41; DB 23; Length 20;
 Best Local Similarity 53.3%; Pred. No. 0.76;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSQSLASV 15
 ||| |||:|:|
 DB 6 DSENSNFRNEIQSLV 20

RESULT 8
 ID ABB81752 standard; protein; 121 AA.

XX ABB81752;

DT 10-SEP-2002 (first entry)

DE Tumour necrosis factor receptor Fas death domain (longer sequence).

XX Tumour necrosis factor; receptor; TNFR-1; death domain;

KW TNFR-1 DD; Fas; protein co-ordinate data.

XX Unidentified.

XX US2002045578-A1.

PD 18-APR-2002.

XX 14-MAY-2001; 2001US-0854906.

XX 22-MAY-2000; 2000US-206215P.

PA (SUKI/) SUKITS S F.

PA (XUGG/) XU G.

PA (LINL/) LIN L.

PA (TELL/) TELLIEZ J.

PA (HSUS/) HSU S.

PI Sukits SF, Xu G, Lin L, Telliez J, Hsu S;

DR WPI; 2002-443412/47.

PT Solution comprising tumor necrosis factor receptor 1 death domain,
 useful for identifying potential inhibitor of tumor necrosis factor

PT receptor 1 death domain -

XX Example 1, Fig 7; 49pp; English.

CC The sequence represents the tumor necrosis factor receptor Fas death
 CC domain (Fas DD). The invention relates to a novel solution comprising a
 CC tumor necrosis factor receptor 1 death domain. The solution is useful
 CC for identifying a potential inhibitor of TNFR-1 DD, for the design and
 CC selection of potent and selective inhibitors of TNF signalling pathways,
 CC and for generating a three-dimensional structure for an unknown molecule
 CC or molecular complex.

XX Sequence 121 AA;

Query Match 56.2%; Score 41; DB 23; Length 121;
 Best Local Similarity 53.3%; Pred. No. 5.9;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSQSLASV 15
 ||| |||:|:|
 DB 105 DSENSNFRNEIQSLV 119

RESULT 9
 ID AAB26982 standard; Protein; 281 AA.

XX AAB26982;

DT 02-FEB-2001 (first entry)

DE Human Fas.

XX Human; Fas; tumour necrosis factor; TR9 receptor; immunosuppressive;
 KW antiinflammatory; cardiant; antidiabetic; antiallergic;
 KW anticholesteric; antirheumatic; anti-HIV; anticonvulsant; cyrostatic;
 KW neuroprotective; gene therapy; Death domain containing Receptor 6;
 KW common variable immunodeficiency; X-linked agammaglobulinemia;
 KW severe combined immunodeficiency; Wiskott-Aldrich syndrome;
 KW autoimmune disease; rheumatoid arthritis; allergic encephalomyelitis;
 KW multiple sclerosis; diabetes mellitus; asthma; epilepsy; cancer;
 KW cardiovascular disease; neurological disease; protein coordinate data.

XX Homo sapiens.

XX WO200056862-A1.

PD 28-SEP-2000.

XX 16-MAR-2000; 2000WO-US06831.

XX 24-MAR-1999; 99US-0126019.

XX 14-MAY-1999; 99US-0134220.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ni J, Gentz RL, Yu G, Fan P;

DR WPI; 2000-594575/56.

XX Nucleic acid molecule encoding a human tumor necrosis factor receptor,
 PT known as TR9, useful for treating, preventing and diagnosing severe
 PT combined immunodeficiency, autoimmune diseases, HIV infection, epilepsy
 PT and cancer -

XX Disclosure, Fig 2, 220pp; English.

CC The present sequence is Fas, a member of the tumor necrosis factor
 CC receptor family. A novel human tumor necrosis factor receptor,
 CC designated TR9, has been isolated. The TR9 receptor is also known as
 CC Death Domain containing Receptor 6. TR9 polypeptides, polynucleotides or
 CC agonists are useful for treating, preventing or diagnosing common
 CC variable immunodeficiency, X-linked agammaglobulinemia, severe combined
 CC immunodeficiency and Wiskott-Aldrich syndrome, autoimmune diseases (such
 CC as rheumatoid arthritis, allergic encephalomyelitis, multiple sclerosis,
 CC diabetes mellitus and asthma), HIV infection, epilepsy, cancer,
 CC cardiovascular diseases and other neurological diseases.

XX Sequence 281 AA;

Query Match 56.2%; Score 41; DB 21; Length 281;
 Best Local Similarity 53.3%; Pred. No. 16;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSQSLASV 15
 ||| |||:|:|
 DB 267 DSENSNFRNEIQSLV 281

RESULT 10
 ID AAR76238 standard; Protein; 314 AA.

XX AAR76238;
 AC AAR76238;
 XX

```

DT 25-MAR-2003 (updated)
DT 06-NOV-1995 (first entry)
XX
DE Fas-delta-TM.
XX
XX Fas-delta-TM; transmembrane deletion; apoptosis; antibody;
XX adoptive immunotherapy; transgenic animal.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..16
FT /label= Sig_peptide
XX
PN WO9513701-A1.
XX
PD 26-MAY-1995.
XX
PF 15-NOV-1994; 94MO-US13173.
XX
PR 15-NOV-1993; 93US-0152443.
XX
PA (LXR-B-) LXR BIOTECHNOLOGY INC.
XX
PI Barr PJ, Kiefer MC, Shapiro JP;
XX
DR WPI: 1995-200120/26.
DR N-PSDB; AAQ93879.
XX
XX New nucleic acid encoding Fas protein without its trans-membrane region
PT - and related vector, transformed cells, transgenic animals, protein and
PT antibodies, useful for control of Fas mediated apoptosis
XX
PS Claim 9; Fig.3-1 to 3-4; 38pp; English.
XX
XX mRNA was obtd. from human lymphocytes and PCR was used to make
CC cDNA specific for Fas-delta-TM (i.e. Fas lacking the transmembrane
CC region) mRNA. The PCR product was ligated into pBluescript and the
CC recombinant plasmid was used to transfect E. coli DH5-alpha cells. The
CC insert sequence of pBluescript-Fas-delta-TM encoded the protein
CC given in AAR76238.
CC (updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 314 AA;
XX
Query Match 56.2%; Score 41; DB 16; Length 314;
Best Local Similarity 53.3%; Pred. No. 18;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1 DSEMYNFRSOLASV 15
DB 300 DSENSNFRNEIOSLV 314
XX
RESULT 11
ID AAR9682 standard; Protein; 314 AA.
XX
AC AAR9682;
XX
DT 10-OCT-1996 (first entry)
XX
DE Human Fas soluble antigen Fas dell.
XX
XX Fas antigen; autoimmune disease; systemic lupus erythematosus; SLE;
XX angiotumoblastic lymphadenopathy; AILD.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..16
FT /label= Sig_peptide
FT Protein 17..314

```

```

FT /label= Mat protein
FT /note= "soluble Fas dell antigen"
FT 17..168
FT Domain
FT /label= Extracellular domain
FT /note= "the 5 C-terminal residues of the
FT Fas antigen extracellular domain are
FT deleted in Fas dell"
FT 169..314
FT Domain
FT /label= Cytoplasmic domain
FT Peptide 164..173
FT /note= "preferred peptide from breakpoint region
FT (claim 4, page 132)"
FT Peptide 164..174
FT /note= "preferred peptide from breakpoint region"
FT Peptide 161..171
FT /note= "preferred peptide from breakpoint region"
XX
XX WO9620206-A1.
XX
XX 04-JUL-1996.
XX
PD 22-DEC-1995; 95MO-US17083.
XX
PR 23-DEC-1994; 94US-0371263.
XX
PA (UABR-) UAB RES FOUND.
XX
XX Cheng J, Liu C, Mountz JD, Zhou T;
XX
XX WPI: 1996-321796/32.
XX
XX N-PSDB; AAT34527.
XX
XX Natural, soluble form of Fas antigen secreted by human cells is
PT result of alternative mRNA processing - used to diagnose
PT Fas-associated disease, e.g. systemic lupus erythematosus
XX
XX Claim 4; Page 114-16; 152pp; English.
XX
XX A natural, soluble Fas antigen variant (AAR9682), designated Fas
CC dell, and other Fas variants (AAR9683-85) are derived by alternative
CC splicing of Fas gene transcripts. A cDNA clone (AAT34527) coding for
CC the variant was obtd. from human peripheral blood mononuclear cells.
CC The Fas dell variant lacks the transmembrane domain of insoluble
CC Fas antigen (AAR9681). Recombinant dell variant, or fragments of
CC it, can be expressed in prokaryotic or eukaryotic (e.g. COS) cells.
CC Detection of increased levels of soluble forms of Fas antigen can
CC be used to diagnose autoimmune diseases, esp. systemic lupus
CC erythematosus and angiotumoblastic lymphadenopathy.
XX
SQ Sequence 314 AA;
XX
Query Match 56.2%; Score 41; DB 17; Length 314;
Best Local Similarity 53.3%; Pred. No. 18;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1 DSEMYNFRSOLASV 15
DB 300 DSENSNFRNEIOSLV 314
XX
RESULT 12
ID AAM98070 standard; Protein; 314 AA.
XX
AC AAM98070;
XX
DT 21-JUN-1999 (first entry)
XX
DE Soluble Fas receptor.
XX
XX Fas receptor; Fas ligand; FasL; proinflammatory; immunosuppressive;
XX graft versus host disease; autoimmune disease; psoriasis;
XX rheumatoid arthritis; systemic lupus erythematosus; gene therapy.
XX

```

```

XX OS Mammalia.
XX FH Key Location/Qualifiers
XX FT Peptide 1..16
XX FT /note= "signal peptide"
XX FT Protein 17..314
XX FT /note= "mature protein"
XX FT Misc-difference 109
XX FT /note= "encoded by GAA"
XX FT Modified-site 118
XX FT /note= "N-glycosylated"
XX PN MO990399-A1.
XX PD 28-JAN-1999.
XX PP 16-JUL-1998; 98WO-US14771.
XX PR 17-JUL-1997; 97US-0052829.
XX PA (UNMI ) UNIV MICHIGAN.
XX PI Chen J, Nabel GJ.
XX DR WPI: 1999-132243/11.
XX DR N-PSDB; AAX24878.
XX PT Inhibition of proinflammatory responses - using an agent which
XX PT modulates FasL stimulation, used for treating graft versus host
XX PT disease or autoimmune disease
XX PS Disclosure; Fig 4B, 71pp; English.
XX CC This present sequence is a soluble Fas receptor. The invention
XX CC provides a method for inhibiting a proinflammatory response in a
XX CC cell mixture by administering an immunosuppressive agent which
XX CC inhibits the proinflammatory activity of Fas ligand (FasL). In some
XX CC embodiments, an FasL is coadministered with the immunosuppressive
XX CC agent, and the cell mixture comprises neutrophil cells. The method
XX CC can be practiced in vitro, ex vivo or in vivo. Suitable
XX CC immunosuppressive agents include antisease molecules that inhibit
XX CC endogenous FasL expression, soluble Fas receptors or variants,
XX CC ribozymes that inhibit the endogenous expression of FasL, drugs
XX CC that inhibit FasL signaling, agents that induce the endogenous
XX CC expression of transforming growth factor (TGF)-beta, and
XX CC polynucleotides coding for an immunosuppressive agent such as
XX CC TGF-beta. The method can be used for treating diseases associated
XX CC with an undesired FasL-mediated proinflammatory response, e.g.
XX CC graft versus host disease, or an autoimmune disease such as
XX CC systemic lupus erythematosus, rheumatoid arthritis and psoriasis.
XX CC The invention also provides a method for identifying agents which
XX CC modulate FasL stimulation of a proinflammatory response.
XX SQ Sequence 314 AA;
QY Query Match 56.2%; Score 41; DB 20; Length 314;
DB Best Local Similarity 53.3%; Pred. No. 18;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 DSEMYNFRSOLASV 15
DB 300 DSENSNFRNEIQLV 314

```

RESULT 13
AAB50893
ID AAB50893 standard; Protein; 331 AA.

AC AAB50893;
XX
DT 19-MAR-2001 (first entry)
XX

```

DE Human Fas receptor.
XX Human, TR10 receptor; cytostatic; immunosuppressive; neuroprotective;
XX antiinflammatory; anti-HIV; antiparkinsonian; neurotropic; cardiant;
XX vasoectopic; antiallergic; antidiabetic; vulnery; ophthalmological;
XX antiviral; antibacterial; antifungal; antiparasitic; gene therapy;
XX tumour necrosis factor receptor; cancer; leukaemia; autoimmune disorder;
XX apoptosis; cardiovascular disorder; inflammatory disease; wound;
XX infection; neurological disease; Fas receptor; protein coordinate data.
XX OS Homo sapiens.
XX PN WO200073321-A1.
XX PD 07-DEC-2000.
XX PP 26-MAY-2000; 2000WO-US14554.
XX PR 28-MAY-1999; 99US-0136786.
XX PR 07-JUL-1999; 98US-0142563.
XX PR 15-JUL-1999; 99US-0144023.
XX PA (HDMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ni J.
XX DR WPI: 2001-025250/03.
XX PT Nucleic acid encoding a tumor necrosis factor receptor 10, useful in
XX PT the diagnosis, treatment or prevention of cancer, autoimmune disorders,
XX PT and diseases and disorders associated with apoptosis -
XX PS Disclosure; Fig 2, 212pp; English.
XX CC The present sequence is given in a specification relating to an isolated
XX CC nucleic acid encoding a human tumour necrosis factor receptor TR10.
XX CC The TR10 polynucleotide, polypeptide, antibodies, agonists and
XX CC antagonists are useful in the diagnosis, treatment or prevention of
XX CC cancer, such as breast and ovarian cancer and leukaemia; autoimmune
XX CC disorders such as multiple sclerosis, Crohn's disease and graft versus
XX CC host disease; diseases associated with increased apoptosis such as AIDS,
XX CC Alzheimer's disease and Parkinson's disease; cardiovascular disorders
XX CC such as limb ischaemia and congenital heart defects; inflammatory
XX CC diseases e.g. allergy; wound healing; disorders associated with
XX CC neovascularisation, e.g. diabetic retinopathy; infectious diseases such
XX CC as viral, bacterial, fungal and parasitic infections; and neurological
XX CC diseases such as amyotrophic lateral sclerosis.
XX SQ Sequence 331 AA;
QY Query Match 56.2%; Score 41; DB 22; Length 331;
DB Best Local Similarity 53.3%; Pred. No. 19;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 DSEMYNFRSOLASV 15
DB 317 DSENSNFRNEIQLV 331

```

RESULT 14
AAR28084
ID AAR28084 standard; Protein; 335 AA.

AC AAR28084;
XX
DT 25-MAR-2003 (updated)
DT 12-MAR-1993 (first entry)
XX
DE Human cell surface antigen.
XX Fas antigen; apoptosis; p58; NGFR/TNFR family.
XX Homo sapiens.
OS

```

XX Key Location/Qualifiers
FH Peptide 1..16
FT Protein /label= signal
FT Protein 17..335
FT Modified-site /label= Fas_antigen
FT Modified-site 118..120
FT /label= N-glycosylation_site
FT /note= "putative"
FT Modified-site 136..138
FT /label= N-glycosylation_site
FT /note= "putative"
FT Domain 174..190
FT /label= transmembrane
FT Domain 17..173
FT /label= extracellular
FT /note= "cysteine-rich"
FT Domain 191..335
FT /label= cytoplasmic
XX EP510691-A1.
XX 28-OCT-1992.
XX 24-APR-1992; 92EP-0107060.
XX 26-APR-1991; 91JP-0125234.
XX (OSAB-) OSAKA BIOSCIENCE INST.
XX PI Itoh N, Nagata S, Yonehara S;
XX DR MPI; 1992-358914/44.
XX N-PSDB; AAQ29959.
XX DNA encoding human cell surface antigen - used to clarify
XX PT apoptosis mechanism of various types of cell, and to prepare
XX PR monoclonal antibodies that react with tumour cells expressing Fas
XX PS Claim 3; Fig 1 and 2; 27pp; English.
XX CC The Fas antigen is implicated in apoptosis. A cDNA clone encoding
XX CC the antigen was isolated (pF58) and the amino acid sequence of Fas
XX CC was deduced from it. The mature protein has a calculated mol.wt. of
XX CC 36,000 and is a member of the NGFR/TNFR family of cell-surface
XX CC membrane proteins. The inventors claim a protein comprising at
XX CC least the extracellular domain of Fas antigen.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 335 AA;

Query Match 56.2%; Score 41; DB 13; Length 335;
Best Local Similarity 53.3%; Pred. No. 19;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSOLASV 15
DB 321 DSENSNFRNEIQSLV 335

```

```

OS Homo sapiens.
XX Key Location/Qualifiers
FH Peptide 1..16
FT Protein /label= sig_peptide
FT Peptide 17..335
FT /label= mat_peptide
XX JP07115988-A.
XX 09-MAY-1995.
XX 26-OCT-1993; 93JP-0267644.
XX 26-OCT-1993; 93JP-0267644.
XX (NISB ) JAPAN TOBACCO INC.
XX PA
XX DR MPI; 1995-202847/27.
XX N-PSDB; AAQ95297.
XX Preparation of soluble membrane proteins - for their use in antibody
XX PT production for the treatment and prevention of related diseases
XX PS Example 1; Pages 15-17; 51pp; Japanese.
XX CC AAR78606 (human Fas protein) is encoded by the plasmid pF58 which
XX CC contains hFas cDNA. The plasmid was used in the construction of an
XX CC expression vector for the prodn. of recombinant soluble membrane
XX CC proteins. The proteins can be used in antibody prodn. for the
XX CC treatment and prevention of related diseases.
XX SQ Sequence 335 AA;

Query Match 56.2%; Score 41; DB 16; Length 335;
Best Local Similarity 53.3%; Pred. No. 19;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSOLASV 15
DB 321 DSENSNFRNEIQSLV 335

Search completed: November 26, 2003, 12:30:27
Job time : 42:5663 secs

```

```

RESULT 15
AAR78606
ID AAR78606 standard; Protein; 335 AA.
XX AAR78606;
XX 19-FEB-1996 (first entry)
XX Human Fas protein.
XX Plasmid pF58; human Fas cDNA; soluble membrane protein;
XX antibody production; diseases; treatment; prevention.
XX

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:22:05 ; Search time 7.22892 Seconds

Title: US-09-230-111C-16

Perfect score: 73
Sequence: 1 DSEMYNFRSGLASV 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database : SW188Prot_41:4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	41	56.2	335	1	TNR6_HUMAN	P25445 homo sapien
2	41	56.2	1066	1	T145_YEAST	P46677 saccharomy
3	39	53.4	995	1	DBPA_YEAST	Q12389 saccharomy
4	38	52.1	152	1	G12_BRARE	P47805 brachydanid
5	38	52.1	252	1	YAG1_ECOLI	P77300 escherichia
6	38	52.1	1709	1	SN_HUMAN	Q9b222 homo sapien
7	38	52.1	2133	1	PA8_PIG	P12633 sus scrofa
8	37	50.7	241	1	Y705_CANJE	Q9pjk2 campylobact
9	37	50.7	255	1	STX8_YEAST	P31377 saccharomy
10	37	50.7	633	1	YN33_YEAST	P48566 saccharomy
11	37	50.7	941	1	CHRD_XENLA	Q91173 xenopus lae
12	37	50.7	1129	1	AK11_RAT	P62924 rattus norv
13	36	49.3	263	1	SPAR_SALTU	P40701 salmonella
14	36	49.3	273	1	VPRT_ASFB7	Q00946 african swi
15	36	49.3	797	1	DPOM_AGABT	P30322 agarticus b
16	36	49.3	4540	1	DTHC_PASTE	Q27111 paramectum
17	35	47.9	191	1	ENGB_MYCE	P47577 mycoplasma
18	35	47.9	302	1	CYSD_BUCAI	P57499 buchnera ap
19	35	47.9	386	1	CMG2_HUMAN	P58335 homo sapien
20	35	47.9	421	1	PHOX_KILTA	P08840 kluyveromy
21	35	47.9	732	1	YMHK_YEAST	Q03254 saccharomy
22	35	47.9	864	1	CHEA_BORBU	Q44737 borrelia bu
23	35	47.9	880	1	RA50_PYEAB	Q9uz08 pyrococcus
24	35	47.9	1885	1	RRPO_ACLSA	P54891 apple chlor
25	34	46.6	66	1	RL29_THBTN	Q87482 thermomane
26	34	46.6	69	1	RL29_CLOPE	Q8xh11 clostridium
27	34	46.6	102	1	INSE_ECOLI	P77661 escherichia
28	34	46.6	102	1	INSE_SHIFL	P59445 shigella fl
29	34	46.6	188	1	RM05_MARPO	P26860 marchantia
30	34	46.6	201	1	TRI_BONLE	P05547 pontastruca
31	34	46.6	237	1	Y450_MYCE	P47688 mycoplasma
32	34	46.6	264	1	DMPH_PESP	P49356 pseudomona
33	34	46.6	264	1	XYLI_PSSPU	P49356 pseudomona

34	34	273	1	VRT	ASF2	Q65228	anabiana swi
35	34	295	1	MIA	ANAP	O8y1n2	anabrena sp
36	34	350	1	UIR8	HSV6	P24438	human herpes
37	34	452	1	KIC0	MOUSE	O61414	mus
38	34	456	1	KIC0	HUMAN	P19012	homo sapien
39	34	466	1	PCPB	SPHR	P42335	sphingobium
40	34	507	1	IF2P	SULSO	O980q8	sulfolobus
41	34	720	1	SPOT	MYCGE	P47520	mycoplasma
42	34	724	1	PALV	YSTMA	O96477	yeastlago ma
43	34	795	1	PHY1	SYN3	O51568	synecchocystis
44	34	848	1	SECA	SYACA	O19911	cyanidium c
45	34	1694	1	SN	MOUSE	Q62230	mus musculu

ALIGNMENTS

ID	TNR6_HUMAN	STANDARD:	PRT:	335 AA
AC	P25445	Q14292; Q14293; Q14294; Q14295; Q16652;		
DT	01-MAY-1992	(Rel. 22, Created)		
DT	01-MAY-1992	(Rel. 22, Last sequence update)		
DT	15-SEP-2003	(Rel. 42, Last annotation update)		
DE	Tumor necrosis factor receptor superfamily member 6 precursor (FASL receptor) (Apoptosis-mediating surface antigen FAS) (Apo-1 antigen) (CD95).			
GN	TNFRSF6 OR APT1 OR FAS OR FASL.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxId=9606;			
RA	Medline=91309137; PubMed=1713127;			
RA	Itch N., Yonehara S., Ishii A., Yonehara M., Mizushima S.I., Sameshima M., Hase A., Seto Y., Nagata S.;			
RT	"The polypeptide encoded by the cDNA for human cell surface antigen Fas can mediate apoptosis.";			
RL	Cell 66:233-243 (1991).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND SEQUENCE OF 226-240; 269-291 AND 321-335.			
RX	MedLINE=92268122; PubMed=1375228;			
RA	Oehm A., Behrmann I., Falk W., Pawlita M., Maier G., Klas C., Li-Weber M., Richards S., Dhein J., Trauth B.C., Ponsingl H., Krammer P.H.;			
RT	"Proliferation and molecular cloning of the Apo-1 cell surface antigen, a member of the tumor necrosis factor/nerve growth factor receptor superfamily. Sequence identity with the Fas antigen.";			
RL	J. Biol. Chem. 267:10709-10715 (1992).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 6), AND FUNCTION.			
RX	MedLINE=95181785; PubMed=7533181;			
RA	Cascino I., Fucci G., Papoff G., Ruberti G.;			
RT	"Three functional soluble forms of the human apoptosis-inducing Fas molecule are produced by alternative splicing.";			
RL	J. Immunol. 154:2706-2713 (1995).			
RN	[4]			
RP	SEQUENCE FROM N.A. (ISOFORM 5).			
RC	TISSUE=Peripheral blood lymphocytes;			
RA	Schaeublein C.B., Boehmann R., Phillipsen P., Babel H.;			
RL	Submitted (UDN-1995) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A. (ISOFORMS 4 AND 5).			
RX	MedLINE=9628926; PubMed=8648105;			
RA	Papoff G., Cascino I., Erano A., Starace G., Lynch D.H., Ruberti G.;			
RT	"An N-terminal domain shared by Fas/Apo-1 (CD95) soluble variance prevents cell death in vitro.";			
RL	J. Immunol. 156:4622-4630 (1996).			
RN	[6]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Urinary bladder;			

RA MEDLINE=2238257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang Y., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stalcup M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Lounellano N.A., Peters G.J., Abramson R.D., Mollahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunnarone P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Maria M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RL [17]
 RP INTERACTION WITH RIPK1.
 RA MEDLINE=95277838; PubMed=7538908;
 RA Stanger B.Z., Leder P., Lee T.-H., Kim E., Seed B.,
 RT "RIP: a novel protein containing a death domain that interacts with
 RT Fas/Apo-1 (CD95) in yeast and causes cell death.";
 RL Cell 81:513-523(1995).
 RL [8]
 RP STRUCTURE BY NMR OF 218-335.
 RA MEDLINE=97122332; PubMed=8967952;
 RA Huang B., Eberstadt M., Olejniczak E.T., Meadows R.P., Feak S.W.,
 RT "NMR structure and mutagenesis of the Fas (APO-1/CD95) death domain.";
 RL Nature 384:638-644(1996).
 RL [9]
 RP VARIANT ALPS PRO-241.
 RA MEDLINE=95300225; PubMed=7540117;
 RA Fisher G.H., Rosenberg F.J., Straus S.E., Dale J.K., Middleton L.A.,
 RA Lin A.Y., Strober W., Lenardo M.J., Puck J.M.,
 RT "Dominant interfering Fas gene mutations impair apoptosis in a human
 RT autoimmune lymphoproliferative syndrome.";
 RL Cell 81:935-946(1995).
 RL [10]
 RP VARIANT ALPS TYR-260.
 RA MEDLINE=97066823; PubMed=8929361;
 RA Drappa J., Vaisnaw A.K., Sullivan K.E., Chu J.-L., Elkon K.B.,
 RT "Fas gene mutations in the Canale-Smith syndrome, an inherited
 RT lymphoproliferative disorder associated with autoimmunity.";
 RL New Engl. J. Med. 335:1643-1649(1996).
 RL [11]
 RP VARIANTS ALPS TRP-121 AND CYS-232.
 RA MEDLINE=97180145; PubMed=9028321;
 RA Bettinelli A., Brugnoni D., Quiroz-Roldan E., Malagoli A.,
 RA La Grutta S., Corrao A., Notarangelo L.D.,
 RT "Missense mutations in the Fas gene resulting in autoimmune
 RT lymphoproliferative syndrome: a molecular and immunological
 RT analysis.";
 RL Blood 89:902-909(1997).
 RL [12]
 RP VARIANTS ALPS ASP-257 AND SER-310.
 RA MEDLINE=97180739; PubMed=9028957;
 RA Sneller M.C., Wang J., Dale J.K., Strober W., Middleton L.A., Choi Y.,
 RA Fleisher T.A., Lim M.S., Jaffe E.S., Puck J.M., Lenardo M.J.,
 RA Straus S.E.,
 RT "Clinical, immunologic, and genetic features of an autoimmune
 RT lymphoproliferative syndrome associated with abnormal lymphocyte
 RT apoptosis.";
 RL Blood 89:1341-1348(1997).
 RL [13]
 RP VARIANT ALPS ALA-28.
 RA MEDLINE=97463633; PubMed=9322534;
 RA Penati L., Costanzo A., Ianni A., Accapezzato D., Iorio R.,
 RA Natoli G., Nisini R., Almerighi C., Balsano C., Vajro P., Vegnente A.,

RA Levero M.,
 RT "Fas/Apo1 mutations and autoimmune lymphoproliferative syndrome in a
 RT patient with type 2 autoimmune hepatitis.";
 RL Gastroenterology 113:1384-1389(1997).
 RL [14]
 RP VARIANT ALPS VAL-260.
 RA MEDLINE=99038860; PubMed=9821419;
 RA Infante A.J., Britton H.A., DeNapoli T., Middleton L.A., Lenardo M.J.,
 RA Jackson C.E., Wang J., Fleisher T.A., Straus S.E., Puck J.M.,
 RT "The clinical spectrum in a large kindred with autoimmune
 RT lymphoproliferative syndrome caused by a Fas mutation that impairs
 RT lymphocyte apoptosis.";
 RL J. Pediatr. 133:629-633(1998).
 RL [15]
 RP VARIANTS ALPS LYS-241 AND GLN-250.
 RA MEDLINE=99192346; PubMed=10090885.
 RA Jackson C.E., Fleisher T.A., Heu A.P., Anderson S.M., Choi Y., Wang J.,
 RA Dale J.K., Fleisher T.A., Middleton L.A., Sneller M.C., Lenardo M.J.,
 RA Straus S.E., Puck J.M.,
 RT "Autoimmune lymphoproliferative syndrome with defective Fas: genotype
 RT influences penetrance.";
 RL Am. J. Hum. Genet. 64:1002-1014(1999).
 RL [16]
 RP VARIANT ALPS GLY-272.
 RA MEDLINE=99270228; PubMed=10340403;
 RA Peters A.M., Kohfink B., Martin H., Griesinger F., Wormann B.,
 RA Gahr M., Roessler J.,
 RT "Defective apoptosis due to a point mutation in the death domain of
 RT CD95 associated with autoimmune lymphoproliferative syndrome, T-cell
 RT lymphoma, and Hodgkin's disease.";
 RL Exp. Hematol. 27:868-874(1999).
 RL [17]
 RP VARIANTS ALPS ARG-82; PRO-250; GLY-260 AND ILE-270.
 RA MEDLINE=99126461; PubMed=9927496;
 RA Vaisnaw A.K., Ofitnick J.R., Chu J.-L., Krammer P.H., Chao M.V.,
 RA Elkon K.B.,
 RT "The molecular basis for apoptotic defects in patients with CD95
 RT (Fas/Apo-1) mutations.";
 RL J. Clin. Invest. 103:355-363(1999).
 RL [18]
 RP VARIANTS NON-HODGKIN'S LYMPHOMA THR-25; PHE-180; LEU-183; ILE-198;
 RP VAL-260; LYS-264; PHE-272 AND ASN-299.
 RA MEDLINE=99005325; PubMed=9787134;
 RA Groenbaek K., Straten P.T., Ralfkiaer E., Ahrekiel V., Andersen M.K.,
 RA Hansen N.E., Zeuben J., Hou-Jensen K., Guldberg P.,
 RT "Somatic Fas mutations in non-Hodgkin's lymphoma: association with
 RT extranodal disease and autoimmunity.";
 RL Blood 92:3018-3024(1998).
 RL [19]
 RP VARIANTS ALPS PRO-241; VAL-260; ILE-270 AND GLY-272.
 RA MEDLINE=2311411; PubMed=11419480;
 RA Straus S.E., Jaffe E.S., Puck J.M., Dale J.K., Elkon K.B.,
 RA Roosen-Wolf A., Peters C.M., Lin A.Y., Beumler C., Siegest E.,
 RA Fisher R.E., Jackson C.M., Lin A.Y., Beumler C., Siegest E.,
 RA Marx A., Vaisnaw A.K., Grodzicky T., Fleisher T.A., Lenardo M.J.,
 RT "The development of lymphomas in families with autoimmune
 RT lymphoproliferative syndrome with germline Fas mutations and
 RT defective lymphocyte apoptosis.";
 RL Blood 98:194-200(2001).
 RL [20]
 RP -1- FUNCTION: Receptor for TNFSF6/FasL. The adaptor molecule FADD
 CC recruits caspase-8 to the activated receptor. The resulting death-
 CC inducing signaling complex (DISC) performs caspase-8 proteolytic
 CC activation which initiates the subsequent cascade of caspases
 CC (aspartate-specific cysteine proteases) mediating apoptosis. FAS-
 CC mediated apoptosis may have a role in the induction of peripheral
 CC tolerance, in the antigen-stimulated suicide of mature T-cells, or
 CC both. The secreted isoforms 2 to 6 block apoptosis (in vitro).
 CC -1- SUBUNIT: Binds DAXX (By similarity). Binds RIPK1.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1);
 CC Secreted (isoforms 2 to 6).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=6;
 CC Name=1;


```

FT NP_BIND 181 188 ATP (POTENTIAL).
FT SITE 288 291 DEAD BOX
SQ SEQUENCE 995 AA; 113157 MW; 9AFA74D0B8AB198A2 CRC64;

Query Match
Best Local Similarity 53.4%; Score 39; DB 1; Length 995;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 3 EMTNFRSQSLASV 15
Db 474 OLYNFRAGLTSL 486

RESULT 4
G12_BRARE STANDARD; PRT; 152 AA.
AC P47805;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gastrulation specific protein G12.
G12.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=96081378; PubMed=8541223;
RA Conway G.;
RT "A novel gene expressed during zebrafish gastrulation identified by
RT differential RNA display."
MEG. Dev. 52:383-391(1995).
CC -1- TISSUE SPECIFICITY: Expressed for a short period in the cells that
CC will produce the enveloping layer.
CC -1- DEVELOPMENTAL STAGE: Gastrulation specific.
CC -1- SIMILARITY: BELONGS TO THE SPOT14 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U27121; AAA96952.1; -
CC ZFIN; ZDB-GENE-990415-81; G12.
DR ZFIN; ZDB-GENE-990415-81; G12.
SQ SEQUENCE 152 AA; 17486 MW; 140BE173FC8B0A2 CRC64;

Query Match
Best Local Similarity 52.1%; Score 38; DB 1; Length 152;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 DSEMTNFRSQSLASV 14
Db 63 EADMVTSYISQLSKI 76

RESULT 5
YAGI_ECOLI STANDARD; PRT; 252 AA.
AC P77300;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical transcriptional regulator yagi.
GN YAGI OR B0272.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=562;

```

```

RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=9742617; PubMed=9278503;
RA Blatner F.R.; Plunkett G. III; Bloch C.A.; Perna N.T.; Burland V.;
RA Riley M.; Collado-Vides J.; Glasner J.D.; Rode C.K.; Mayhew G.F.;
RA Gregor J.; Davis N.W.; Kirkpatrick H.A.; Goeden M.A.; Rose D.J.;
RA Mau B.; Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RN Science 277:1453-1474(1997).
RN (2)
RP SEQUENCE FROM N.A.
RA Schramm S.; Duncan M.; Allen E.; Araujo R.; Aparicio A.; Chung E.;
RA Davis K.; Federspiel N.; Hyman R.; Kallman S.; Komp C.; Kurdi O.;
RA Lashkari D.; Lew H.; Lin D.; Namath A.; Oefner P.; Roberts D.;
RA Davis R.W.;
RT Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ICLR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000135; AAC73375.1; -
CC EMBL; U70214; AB08693.1; -
DR PIR; H64752; H64752.
DR Ecogene; EG13349; YAGI.
DR InterPro; IPR005471; HTH_ICLR.
DR InterPro; IPR005473; HTH_ICLRlike.
DR Pfam; PF01614; ICLR_1.
DR SMART; SM00346; HTH_ICLR; 1.
DR PROSITE; PS01051; HTH_ICLR_FAMILY; 1.
KW Hypothetical protein; Transcription regulation; DNA-binding;
KW Complete proteome.
FT NP_BIND 25 45 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 252 AA; 27838 MW; 8FF297AEB8FF2E435 CRC64;

Query Match
Best Local Similarity 52.1%; Score 38; DB 1; Length 252;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DSEMTNFRSQSL 11
Db 227 DAELANFREQL 237

RESULT 6
SN_HUMAN STANDARD; PRT; 1709 AA.
ID SN_HUMAN
AC Q9BZ22; Q96DL4; Q9GZ55; Q9H1H6; Q9H1H7; Q9H7L7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sialoadhesin precursor (Sialic acid binding Ig-like lectin-1) (Siglec-
DE 1) (CD169 antigen).
GN SN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.
RX TISSUE=Monocytes; PubMed=11133773;
RX MEDLINE=20575418; PubMed=11133773;
RA Hartnell A.; Steel J.; Turley H.; Jones M.; Jackson D.G.;
RA Crocker P.R.;
RT "Characterization of human sialoadhesin, a sialic acid binding
RT receptor expressed by resident and inflammatory macrophage

```

RT populations." ;
 RN (2)
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX MEDLINE=21638749; PubMed=11780053;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 Jones M., Strevrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 Bailey V., Bartlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 R Buck D., Burrill M.D., Butler A.P., Carder C., Carter N.P.,
 Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 R Clegg S., Codley V.E., Collier R.E., Connor R.E., Corby N.R.,
 R Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 R Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 R Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 R Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 R Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 R Kay M.P., Kimberley A.M., King A., Knighte A., Laird G.K., Lawlor S.,
 R Lervaeaelahlo M.H., Levestra M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 R Marsh V.L., Martin S.L., McConachie L.J., Melay K., Murray A.A.,
 R Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 R Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.C.,
 R Phillimore B.J.C.T., Patrathalingam S.R., Plumb R.W., Ramsay H.,
 R Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
 R Skuse C.D., Smith M.L., Sodehlund C., Steward C.A., Sulston J.E.,
 R Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 R Tracey A., Tromans A.C., Vaudin M., Walli M., Wallis J.M.,
 R Whitehead S.L., Whitlaker P., Willey D.L., Williams L., Williams S.A.,
 R Wilmfing L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 R Rogers J.,
 RA "The DNA sequence and comparative analysis of human chromosome 20." ;
 RL Nature 414:865-871(2001).
 RN (3)
 RP SEQUENCE OF 733-1709 FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Spleen;
 RX MEDLINE=21082933; PubMed=11214971;
 RA Hattori A., Okumura K., Nagase T., Kikuno R., Hirogawa M., Ohara O.,
 RL "Characterization of long cDNA clones from human adult spleen." ;
 RT DNA Res. 7:357-366(2000).
 RN (4)
 RP SEQUENCE OF 1539-1709 FROM N.A. (ISOFORM 3).
 RC TISSUE=Thymus;
 RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,
 RA Matsunoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,
 RA Matsunuma Y., Moriya S., Chiba S., Momiyama H., Onogawa S.,
 RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Takasaka R.,
 RA Kuga N., Kurida A., Satoh I., Kamata K., Takami S., Terashima Y.,
 RA Wakematsu A., Ithil S., Yamamoto J., Isono Y., Sato H., Oka T.,
 RA Nishikawa T., Kimura K., Yamashita H., Matsuo K., Nakamura Y.,
 RA Sekine M., Kikuchi H., Kanda K., Nagasuna M., Murakawa K.,
 RA Kanehori K., Takahashi-Pujil A., Oshima A., Sugiyama A., Kawakami B.,
 RA Suzuki Y., Sugano S., Nagahata K., Maehuo Y., Nagai K., Isogai T.,
 RT Submitted (OCT-2001) to the EMBL/Genbank/DBA databases.
 RL -1- FUNCTION: Macrophage-restricted adhesion molecule that mediates
 cell-cell dependent binding to lymphocytes, including
 granulocytes, monocytes, natural killer cells, B-cells and CD8 T-
 cells. Preferentially binds to alpha2,3-linked sialic acid (By
 similarity). Binds to SPN/CD43 on T-cells (By similarity). May
 play a role in hemopoiesis.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1) and
 CC soluble (isoform 2).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Comment=Additional isoforms seem to exist;
 CC Name=1;
 CC IsoId=Q9BZ22-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9BZ22-2; Sequence=VSP_002571;
 CC Name=3;
 CC IsoId=Q9BZ22-3; Sequence=VSP_002572;
 CC Note=No experimental confirmation available;

CC	-1	TISSUE SPECIFICITY: Expressed by macrophages in various tissues.
CC		High levels are found in spleen, lymph node, perivascular
CC		macrophages in brain and lower levels in bone marrow, liver
CC		Kupffer cells and lamina propria of colon and lung. Also expressed
CC		by inflammatory macrophages in rheumatoid arthritis.
CC	-1	SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. SIGLEC
CC		(SIALIC ACID BINDING IG-LIKE LECTIN) SUBFAMILY.
CC	-1	SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC	-1	SIMILARITY: Contains 16 immunoglobulin-like C2-type domains.
CC	-1	DATABASE: NAME=PROW, NOTE=PROW 2.18-22(2001).
CC		WWW="http://www.ncbi.nlm.nih.gov/prov/guide/985165905.g.htm".
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration
CC		between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC		the European Bioinformatics Institute. There are no restrictions on its
CC		use by non-profit institutions as long as its content is in no way
CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (See http://www.isb-eb.ch/announce/
CC		or send an email to license@isb-eb.ch).
CC		-----
DR	EMBL	AF230073; AKK0757.1; -
DR	EMBL	AL109804; CAC17543.1; -
DR	EMBL	AL109804; CAC17543.1; -
DR	EMBL	AK024462; BAB15752.1; -
DR	EMBL	AK024459; BAB15749.1; -
DR	EMBL	AK024479; BAB15769.1; -
DR	EMBL	AK057560; BAB15757.1; -
DR	HSSP	O62230; IQFO.
DR	Genew	HGNC:1127; SN.
DR	GO	GO:0016021; C:integral to membrane; NAS.
DR	GO	GO:0005530; F:lectin; NAS.
DR	GO	GO:0016137; P:cell-cell adhesion; NAS.
DR	GO	GO:0007160; P:cell-matrix adhesion; NAS.
DR	GO	GO:0006954; P:inflammatory response; NAS.
DR	InterPro	IPR007110; IG-like.
DR	InterPro	IPR003598; IG_c2.
DR	InterPro	IPR003006; IG_MHC.
DR	Pfam	PF00047; IG; 14.
DR	SMART	SM00408; IGc2; 4.
DR	PROSITE	PS50835; IG_LIKE; 14.
KW	Cell adhesion; Lectin; Antigen; Transmembrane; Signal; Glycoprotein;	
KW	Immunoglobulin domain; Repeat; Alternative splicing; Polymorphism.	
KW		BY SIMILARITY.
FT	SIGNAL	1 19
FT	CHAIN	20 1709
FT	DOMAIN	20 1641
FT	TRANSSEM	1642 1662
FT	DOMAIN	1663 1709
FT	DOMAIN	20 136
FT	DOMAIN	139 233
FT	DOMAIN	238 320
FT	DOMAIN	326 405
FT	DOMAIN	411 507
FT	DOMAIN	511 593
FT	DOMAIN	601 705
FT	DOMAIN	708 785
FT	DOMAIN	799 894
FT	DOMAIN	898 977
FT	DOMAIN	984 1083
FT	DOMAIN	1085 1165
FT	DOMAIN	1176 1248
FT	DOMAIN	1259 1341
FT	DOMAIN	1350 1442
FT	DOMAIN	1445 1528
FT	DOMAIN	1536 1631
FT	DISULFID	36 166
FT	DISULFID	41 98
FT	DISULFID	160 217
FT	DISULFID	262 305
FT	DISULFID	346 390
FT	DISULFID	433 491
FT	DISULFID	531 575
FT	DISULFID	624 689
FT		BY SIMILARITY.


```

FT CARBOHYD 1919 1919 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CONFLICT 713 713 I -> M (IN REF. 2).
FT CONFLICT 724 724 I -> Q (IN REF. 2).
FT CONFLICT 792 792 G -> T (IN REF. 2).
FT CONFLICT 1133 1133 E -> F (IN REF. 2).
FT CONFLICT 1191 1191 I -> L (IN REF. 2).
FT CONFLICT 1209 1209 R -> F (IN REF. 2).
FT CONFLICT 1437 1437 C -> G (IN REF. 2).
FT CONFLICT 1456 1456 F -> R (IN REF. 2).
FT CONFLICT 1539 1539 F -> R (IN REF. 2).
FT CONFLICT 1546 1546 Q -> N (IN REF. 2).
SQ SEQUENCE 2133 AA; 239304 MW; 152BBA8997F570DA CRC64;

Query Match 52.1%; Score 38; DB 1; Length 2133;
Best Local Similarity 70.0%; Pred. No. 85;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 NFRSOLASV 15
||| ||| |||
Db 916 NFRSOLGATV 925

RESULT 8
Y705_CAMJ2 STANDARD; PRT; 241 AA.
Q9PRK2;
16-OCT-2001 (Rel. 40; Created)
16-OCT-2001 (Rel. 40; Last sequence update)
16-OCT-2001 (Rel. 40; Last annotation update)
DE Hypothetical protein Cj0705.
GN Cj0705
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxId=197;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=NCIC 11168;
RX MEDLINE=20150912; PubMed=10689204;
RA Parkhill J., Wren B.W., Mungall K., Kelsey J.M., Churcher C.,
RA Baeham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagers K., Kariyeh A.V., Moule S., Pallen C.W., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrett B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:655-668(2000).
CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AL139076; CAB72979.1; -.
CC PIR, D81341; D81341.
CC InterPro: IPR002678; DUF34.
CC Pfam: PF01784; NIF3; 1.
CC TIGRfam: TIGR00486; TIGR00486; 1.
CC DR Hypothetical protein; Complete proteome.
CC KW SEQUENCE 241 AA; 27786 MW; 1574F24905B89EA CRC64;

Query Match 50.7%; Score 37; DB 1; Length 241;
Best Local Similarity 63.6%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SEMYNFRSOLA 12
||| ||| |||
Db 4 SEIYNFLDQLS 14

```

```

RESULT 9
ID STX8_YEAST STANDARD; PRT; 255 AA.
AC P31377;
DT 01-JUL-1993 (Rel. 26; Created)
DT 15-SEP-2003 (Rel. 42; Last sequence update)
DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE Syntaxin 8 (SNARE protein related to mammalian syntaxin 8) (ULP1
DE interacting protein 2).
DE SYN8 OR ULP2 OR YAL014C OR FUN34.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetaceae; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxId=4932;
RN [1]
RS SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND INTERACTION
RC WITH PEPI2; VTII AND SNC.
RC STRAIN=BY4742;
RX MEDLINE=22342875; PubMed=12453154;
RA Lewis M.J., Pelham H.R.B.;
RT "A new yeast endosomal SNARE related to mammalian syntaxin 8."
RT Traffic 3:922-929(2002).
RN [2]
RS SEQUENCE FROM N.A.
RC STRAIN=S288c;
RA Burri U., Holmann K., Lithgow T.;
RT "Re-sequencing of the ULP2 gene reveals a frame-correction and a new
RT QC-SNARE in the endosome of Saccharomyces cerevisiae."
RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RS SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=93209532; PubMed=8458570;
RA Ouellette B.F., Clark M.W., Keng T., Storms R.K., Zhong W.W.,
RA Zeng B., Fortin N., Delaney S., Barton A.B., Kaback D.B., Buessey H.;
RT "Sequencing of chromosome I from Saccharomyces cerevisiae: analysis
RT of a 32 kb region between the LTR1 and SP07 genes."
RL Genome 36:32-42(1993).
RN [4]
RS SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=94193531; PubMed=8144453;
RA Barton A.B., Kaback D.B.;
RT "Molecular cloning of chromosome I DNA from Saccharomyces cerevisiae:
RT analysis of the genes in the FUN38-MAK16-SP07 region."
RL J. Bacteriol. 176:1872-1880(1994).
RN [5]
RS REVISIONS TO C-TERMINUS.
RA Dolinski K.J., Cherry J.M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: t-SNARE which may play a role in determining the
CC specificity of membrane fusion, protein transport and vesicle
CC trafficking within the Golgi/endosomal and plasma
CC membrane/endosomal systems.
CC -1- SUBUNIT: Interacts with PEPI2, VTII and the SNC SNARE complex
CC proteins.
CC -1- SUBCELLULAR LOCATION: Type IV membrane protein; endosome.
CC -1- SIMILARITY: BELONGS TO THE SYNTAXIN/EPIMORPHIN FAMILY.
CC -1- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AY205237; AA046883.1; -.
CC DR EMBL, AY205236; AA022142.1; -.
CC DR EMBL, L05146; AAC04943.2; -.
CC PIR, S36720; S36720.

```

DR SGD; S0000012; YAL014C.
 DR InterPro: IPR000727; T SNARE.
 DR PROSITE: P550192; T SNARE; 1.
 KW Transport; Protein Transport; Transmembrane; Coiled coil.
 FT DOMAIN 1 234 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 235 254 ANCHOR FOR TYPE IV MEMBRANE PROTEIN
 FT (POTENTIAL).
 FT T-SNARE COILED-COIL HOMOMOLOGY.
 SQ SEQUENCE 255 AA; 29048 MW; 7631F8672F58E5D6 CRC64;

Query Match 50.7%; Score 37; DB 1; Length 255;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SEMNFRSOLA 12
 Db 76 DKELYRFOQOVA 87

RESULT 10
 YN3 YEAST STANDARD; PRT; 633 AA.
 AC P48566;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical 73.0 kDa protein in CLA4-PUS4 intergenic region.
 GN YNL293W OR N0470.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / FY1679;
 RX MEDLINE=96132033; PubMed=8553702;
 RA Maurer K.C.T.; Urbanus J.H.M.; Planta R.J.;
 RT "Sequence analysis of a 30 kb DNA segment from yeast chromosome XIV
 RT carrying a ribosomal protein gene cluster, the genes encoding a
 RT plasma membrane protein and a subunit of replication factor C, and a
 RT novel putative serine/threonine protein kinase gene."
 RL Yeast 11:1303-1310(1995).
 CC -1- SIMILARITY: Contains 1 Rab-GAP TBC domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; U23084; AAC9106.1; -
 DR EMBL; 271569; CAAG6211.1; -
 DR PIR; S60408; S60408.
 DR SGD; S0005237; MSB3.
 DR GO; GO:0005934; C:bud tip; IDA.
 DR GO; GO:0000131; C:incipient bud site; IDA.
 DR GO; GO:0005097; F:RAB GTPase activator activity; IDA.
 DR InterPro: IPR000195; RadGAP_TBC.
 DR Pfam; PF00566; TBC; 1.
 DR SMART; SM00164; TBC; 1.
 DR PROSITE; PS50086; TBC_RABGAP; 1.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 329 349 POTENTIAL.
 FT TRANSMEM 422 442 POTENTIAL.
 FT DOMAIN 223 456 RAB-GAP TBC.
 SQ SEQUENCE 633 AA; 72999 MW; 5F6607B95BBAFB9 CRC64;

Query Match 50.7%; Score 37; DB 1; Length 633;
 Best Local Similarity 53.8%; Pred. No. 33;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SEMNFRSOLASV 14
 Db 600 SEVGFKSLAGV 612

RESULT 11
 CHRD_XENLA STANDARD; PRT; 941 AA.
 ID CHRD_XENLA
 AC 091713;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chordin precursor (Organizer-specific secreted dorsoalizing factor).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Dorsal lip;
 RX MEDLINE=95094250; PubMed=8001117;
 RA Saasai Y., Lu B., Steinbeisser H., Geisler D., Gont L.K.,
 RA de Robertis E.M.;
 RT "Xenopus chordin: a novel dorsoalizing factor activated by organizer-
 RT specific homeobox genes."
 RL Cell 79:779-790(1994).
 CC -1- FUNCTION: POTENT DORSALIZING FACTOR. HAS POTENT AXIS-FORMING
 CC ACTIVITIES INCLUDING THE ABILITY TO RECRUIT NEIGHBORING CELLS INTO
 CC SECONDARY AXES. REGULATES CELL-CELL INTERACTIONS IN THE ORGANIZING
 CC CENTERS OF HEAD, TRUNK AND TAIL DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: PRECORDAL PLATE, NOTOCHORD, CHORDONEURAL
 CC HINGE, EXPRESSION CLOSELY FOLLOWS THE AREAS OF THE EMBRYO THAT
 CC EXPRESS GOOSECOID AND XNOT2.
 CC -1- DEVELOPMENTAL STAGE: EMBRYOGENESIS.
 CC -1- SIMILARITY: BELONGS TO THE CHORDIN FAMILY.
 CC -1- SIMILARITY: Contains 4 VWF domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; J35764; AAC42222.1; -
 DR PIR; A5195; A5195.
 DR InterPro: IPR006559; SOG.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam; PF00093; VWF; 4.
 DR SMART; SM00566; SOG; 3.
 DR SMART; SM00214; VWF; 4.
 DR PROSITE; PS01208; VWF_C_1; 2.
 DR PROSITE; PS50184; VWF_C_2; 3.
 KW Developmental protein; Repeat; Glycoprotein; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 941 CHORDIN.
 FT DOMAIN 41 117 VWF_C 1.
 FT DOMAIN 691 751 VWF_C 2.
 FT DOMAIN 769 838 VWF_C 3.
 FT DOMAIN 857 919 VWF_C 4.
 FT CARBOHYD 244 244 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 433 433 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 735 735 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 941 AA; 104947 MW; 7D1C3262064C7323 CRC64;

Query Match 50.7%; Score 37; DB 1; Length 941;
 Best Local Similarity 53.8%; Pred. No. 52;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;


```

RP SEQUENCE FROM N.A.
RX MEDLINE=92263807; PubMed=1316688;
RA Garcia-Beato R., Freije J.M., Lopez-Ortin C., Blasco R., Vinuela E.;
RT "A gene homologous to topoisomerase II in African swine fever virus.";
RL Virology 188:938-947(1992).
RN
RP COMPLETE GENOME.
RA Yanez R.J., Rodriguez J.M., Nogal M.L., Yuste L., Enriquez C.,
RA Rodriguez J.F., Vinuela E.;
RT "Analysis of the complete nucleotide sequence of African swine fever
RT virus.";
RL Virology 208:249-278(1995).
RN
RP FUNCTION, CHARACTERIZATION, AND MUTAGENESIS.
RX MEDLINE=20576431; PubMed=11031264;
RA Andres G., Alejo A., Simon-Mateo C., Salas M.L.;
RT "African swine fever virus protease, a new viral member of the
RT SUMO-1-specific protease family.";
RL J. Biol. Chem. 276:780-787(2001).
CC
CC -1- FUNCTION: Protease involved in the processing of the virus
CC POLYPROTEINS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to peptidase family C48.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M88699; AAA42734.1; -
DR EMBL; U18466; AAA65340.1; -
DR PIR; A42549; A42549.
DR MEROPS; C48.050; -
KW Hydrolyase; Thiol protease; Late protein.
FT ACT_SITE 168 168 PROBABLE.
FT ACT_SITE 167 187 POTENTIAL.
FT ACT_SITE 232 232 PROBABLE.
FT MUTAGEN 168 168 H->R: LOSS OF ACTIVITY.
FT MUTAGEN 232 232 C->S: LOSS OF ACTIVITY.
SQ SEQUENCE 273 AA; 31550 MW; C5AA78179C57C17 CRC64;

Query Match 49.3%; Score 36; DB 1; Length 273;
Best Local Similarity 54.5%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSOL 11
| : ||| :|
Db 259 DEDMYKFRTHL 269

RESULT 15
DPOM AGABT STANDARD; PRT; 797 AA.
AC P30322:
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable DNA polymerase (EC 2.7.7.7) (Fragment).
OS Agaricus bisporus (Pavement mushroom).
OG Mitochondrion.
OG Plasmid pEM.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Agaricaceae; Agaricus.
OX NCBI TaxID=5343;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 24666 / AG4;
RX MEDLINE=91347410; PubMed=1879001;
RA Robison M.M., Royer J.C., Horgen P.A.;
RT "Homology between mitochondrial DNA of Agaricus bisporus and an

```

```

RT Internal portion of a linear mitochondrial plasmid of Agaricus
RT bisporus";
RL Curr. Genet. 19:495-502(1991).
CC
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA) (N).
CC -1- MISCELLANEOUS: THIS DNA POLYMERASE REQUIRES A PROTEIN AS A PRIMER
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY. STRONG, TO
CC DNA POLYMERASE OF OTHER FUNGAL AND PLANTS MITOCHONDRIAL PLASMIDS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X63075; CAA44800.1; -
DR PIR; S28103; S28103.
DR InterPro; IPR006172; DNA_pol_B.
DR InterPro; IPR004668; DNA_pol_B_2.
DR Pfam; PF03175; DNA_pol_B_2; 1.
DR SMART; SM00486; POLB; 1.
DR PROSITE; PS00116; DNA_POLYMERASE_B; FALSE NEG.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Plasmid; Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 797 AA; 91922 MW; 5C49EAF51B7F927 CRC64;

Query Match 49.3%; Score 36; DB 1; Length 797;
Best Local Similarity 53.8%; Pred. No. 67;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSOL 13
| : ||| :|
Db 22 DVQMYNFTKLES 34

Search completed: November 26, 2003, 12:31:22
Job time : 9.22892 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:25:40 ; Search time 13.9157 Seconds
(without alignments)
103.662 Million cell updates/sec

Title: US-09-230-111C-16
Perfect score: 73
Sequence: 1 DSEMNFRSQLASV 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	50	68.5	216	2	T38520	hypothetical conse
2	42	57.5	199	2	T49486	hypothetical prote
3	41	56.2	314	2	I37383	FAS soluble protei
4	41	56.2	335	2	A40036	apoptosis-mediati
5	41	56.2	804	2	T14762	hypothetical prote
6	41	56.2	1066	2	S50237	TATA box-binding p
7	40	54.8	356	2	D90313	conserved hypochet
8	40	54.8	712	2	T02479	potassium transpor
9	40	54.8	993	2	F97717	hypothetical prote
10	39	53.4	1124	2	F71719	hypothetical prote
11	39	53.4	492	2	AH1084	hypothetical prote
12	39	53.4	995	2	S67564	hypothetical prote
13	39	53.4	1030	2	S57380	probable membrane
14	38	52.1	252	2	H64752	probable transcrip
15	38	52.1	312	2	T16257	hypothetical prote
16	38	52.1	869	2	A25945	coagulation factor
17	38	52.1	2133	2	T43763	coagulation factor
18	37	50.7	205	2	S36720	FUN34 protein - ye
19	37	50.7	241	2	D81341	hypothetical prote
20	37	50.7	633	2	S60408	probable membrane
21	37	50.7	941	1	A55195	chordin precursor
22	37	50.7	1129	2	T42732	A-kinase anchoring
23	37	50.7	1667	2	T15863	hypothetical prote
24	36.5	50.0	4589	2	T14914	ABC transpore, A
25	36	49.3	234	2	C97001	secretory protein - Sal
26	36	49.3	263	2	AD0851	hypothetical prote
27	36	49.3	273	2	A42549	hypothetical prote
28	36	49.3	282	2	G97709	hypothetical prote
29	36	49.3	282	2	G97709	hypothetical prote

30	36	49.3	376	2	B90391	nifs protein homol
31	36	49.3	484	2	F98067	beta-fructofuranos
32	36	49.3	621	2	T39204	hypothetical prote
33	36	49.3	753	2	H90124	hypothetical prote
34	36	49.3	795	2	T20609	hypothetical prote
35	36	49.3	797	2	S28103	probable DNA-direc
36	36	49.3	4540	2	T10838	cytoplasmic dynein
37	35	47.9	80	2	A86762	hypothetical prote
38	35	47.9	191	1	A64237	hypothetical prote
39	35	47.9	236	2	C81253	probable peptide A
40	35	47.9	286	2	E97241	hypothetical prote
41	35	47.9	302	2	B84979	sulfate adenylylitr
42	35	47.9	333	2	H89961	hypothetical prote
43	35	47.9	374	2	D83328	probable aminotran
44	35	47.9	377	2	T05354	hypothetical prote
45	35	47.9	390	2	AB2381	type I site-specif

ALIGNMENTS

```

RESULT 1
T38520
hypothetical conserved protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T38520
R:Oliver, K.; Harris, D.; Barrett, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: 221748
A:Accession: T38520
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-216 <OLI>
A:Cross-references: EMBL:299259; PIDN:CA816369.1; GSPDB:GN00066; SPDB:SPAC2C4.09
A:Experimental source: strain 972h-, cosmid c2C4
C:Genetics:
A:Gene: SPDB:SPAC2C4.09
A:Map position: 1
A:Introns: 74/2; 203/2

Query Match
Best Local Similarity 68.5%; Score 50; DB 2; Length 216;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSEMNFRSQLAS 13
DB 176 DSEMNFRSQLAS 188

RESULT 2
T49486
hypothetical protein B14D6.380 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
C:Accession: T49486
R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: 225022
A:Accession: T49486
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-199 <SCH>
A:Cross-references: EMBL:ALJ56173; GSPDB:GN00116; NCSP:B14D6.380
A:Experimental source: BAC clone B14D6; strain OR74A
C:Genetics:
A:Gene: NCSP:B14D6.380
A:Map position: 6
C:Superfamily: Neurospora crassa hypothetical protein B14D6.380

Query Match
Best Local Similarity 57.5%; Score 42; DB 2; Length 199;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY 4 MYNRSQSLASV 15
|||:|||||:
Db 1 MYNRSQSLISIV 12

RESULT 3
137383
FAS soluble protein - human
C/Species: Homo sapiens (man)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jul-2000
C/Accession: I37383
R/Cascino, I.; Ficuci, G.; Papoff, G.; Rubertl, G.
J. Immunol. 154, 2706-2713, 1995
A/Title: Three functional soluble forms of the human apoptosis-inducing Fas molecule are
A/Reference number: I37383; MUID:95181785; PMID:7533181
A/Accession: I37383
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-314 <RES>
A/Cross-references: EMBL:247993; NID:9728578; PIDN:CAA88031.1; PID:9695539

Query Match 56.2%; Score 41; DB 2; Length 314;
Best Local Similarity 53.3%; Pred. No. 5.6;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSQSLASV 15
|||:|||||:
Db 300 DSEMYNFRNEIQSLV 314

RESULT 4
A40036
apoptosis-mediating surface antigen Fas precursor - human
N/Alternate names: surface antigen APO-1
C/Species: Homo sapiens (man)
C/Date: 17-Jan-1992 #sequence_revision 17-Jan-1992 #text_change 21-Jul-2000
C/Accession: A40036; S24543; A38142
R/Itob, N.; Yonehara, S.; Ishii, A.; Yonehara, M.; Mizushima, S.I.; Samehima, M.; Hase,
Cell 66, 233-243, 1991
A/Title: The polypeptide encoded by the cDNA for human cell surface antigen Fas can medi
A/Reference number: A40036; MUID:91309137; PMID:1713127
A/Accession: A40036
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-335 <ITO>
A/Cross-references: GB:M67454; NID:9182409; PIDN:AAA63174.1; PID:9182410
R/Kramer, P.H.
submitted to the EMBL Data Library, February 1992
A/Reference number: S24543
A/Accession: S24543
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-335 <KRA>
A/Cross-references: EMBL:X63717; NID:928741; PID:928742
R/Oehm, A.; Behrmann, I.; Falk, W.; Pawlita, M.; Maier, G.; Klas, C.; Li-Weber, M.; Rich
J. Biol. Chem. 267, 10709-10715, 1992
A/Title: Purification and molecular cloning of the APO-1 cell surface antigen, a member
A/Reference number: A38142; MUID:92268122; PMID:1375228
A/Accession: A38142
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: nucleic acid
A/Residues: 1-134, 'Q', 136-335 <OEH>
A/Experimental source: SKW6.4 cells
A/Note: sequence extracted from NCBI backbone (NCBI:103810)
C/Genetics: GDB: APT1
A/Cross-references: GDB:132671; OMIM:134637
A/Map position: 10q24.1-10q24.1
C/Suprafamily: NGF receptor repeat homology
C/Keywords: apoptosis; surface antigen; transmembrane protein
P.1-16/Domain: signal sequence #status predicted <SIG>

F.85-128/Domain: NGF receptor repeat homology <NGA>
F.174-190/Domain: transmembrane #status predicted <TMM>

Query Match 56.2%; Score 41; DB 2; Length 335;
Best Local Similarity 53.3%; Pred. No. 6;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSEMYNFRSQSLASV 15
|||:|||||:
Db 321 DSEMYNFRNEIQSLV 335

RESULT 5
T14762
hypothetical protein DKFP434A014.1 - human. (fragment)
C/Species: Homo sapiens (man)
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C/Accession: T14762
R/Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A/Reference number: T18181
A/Accession: T14762
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-804 <WAM>
A/Cross-references: EMBL:AL110224
C/Genetics: GDB: APT1
A/Note: DKFP434A014.1

Query Match 56.2%; Score 41; DB 2; Length 804;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DSEMYNFRSQSLA 12
|||:|||||:
Db 636 DSEMYNFRSSVA 647

RESULT 6
S50237
TATA box-binding protein-protein-associated factor chain TAFII145 - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein G9374; protein YGR274c; TAFII130 protein
C/Species: Saccharomyces cerevisiae
C/Date: 27-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 21-Jul-2000
C/Accession: S50237; S62389; S62397; S62394; S62395; S64607
R/Reese, J.C.; Apone, L.; Walker, S.S.; Griffin, L.A.; Green, M.R.
submitted to the EMBL Data Library, September 1994
A/Description: Yeast TAFIIs in a multisubunit complex required for activated transcripti
A/Reference number: S50237
A/Accession: S50237
A/Molecule type: DNA
A/Residues: 1-1066 <REE>
A/Cross-references: EMBL:U14954; NID:9595837; PIDN:AAA79178.1; PID:9595838
R/Reese, J.C.; Apone, L.; Walker, S.S.; Griffin, L.A.; Green, M.R.
Nature 371, 523-527, 1994
A/Title: Yeast TAFIIs in a multisubunit complex required for activated transcription.
A/Reference number: S62389; MUID:95021883; PMID:7935765
A/Accession: S62389
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-1066 <REW>
A/Cross-references: EMBL:U14954; NID:9595837; PIDN:AAA79178.1; PID:9595838
A/Accession: S62397
A/Molecule type: protein
A/Residues: 583-599; 651-671 <REF>
R/Poon, D.; Bai, Y.; Campbell, A.M.; Bjorklund, S.; Kim, Y.J.; Zhou, S.; Kornberg, R.D.;
Proc. Natl. Acad. Sci. U.S.A. 92, 8224-8228, 1995
A/Title: Identification and characterization of a TFIID-like multiprotein complex from S
A/Reference number: S62390; MUID:95396770; PMID:7667272
A/Accession: S62394
A/Molecule type: DNA
A/Residues: 370-830 <POO>

A:Accession: 562395
A:Molecule type: Protein
A:Residues: 368-384/528-554/752-783 <POM>
A:Parzer, L.; Agostoni Carbone, M.L.; Melchiorretto, P.; Plevani, P.; Martegani, E.; Var
submitted to the Protein Sequence Database, May 1996
A:Reference number: 564591
A:Accession: 564607
A:Molecule type: DNA
A:Residues: 1-1066 <PAN>
A:Cross-references: EMBL:Z73059; NID:g1323498; PID:e243586; PID:g1323499; MIPS:YGR274C
A:Experimental source: strain 528C
C:Genetic:
A:Gene: SGD:TAF145; TAF130
A:Cross-references: SGD:S0003506; MIPS:YGR274C
A:Map position: 7K

Query Match 56.2%; Score 41; DB 2; Length 1066;
Best Local Similarity 57.1%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DSEMYNFRSQLSV 14
|||:||||:
Db 774 DNEAYNFDKSLKSL 787

RESULT 7
D90313
conserved hypothetical protein [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C:Accession: D90313
R:Shu, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweez, M.U.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
arrest, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A89139
A:Accession: D90313
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-356 <KUR>
C:Cross-references: GB:AE006641; NID:g13814777; PIDN:AAK41763.1; GSPDB:GN00155
C:Genetic8:
A:Gene: SSO1545
C:Superfamily: Pyrococcus abyssi hypothetical protein PAB1618

Query Match 54.8%; Score 40; DB 2; Length 356;
Best Local Similarity 63.6%; Pred. No. 10;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DSEMYNFRSQL 11
|||:||||:
Db 321 DSEIYNVLTQL 331

RESULT 8
T02479
potassium transport protein KUP1, high-affinity - Arabidopsis thaliana
M:Alternate names: protein T27B13.19
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 23-Mar-2001
C:Accession: T00591; T02479; B84704
R:Kounisley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL data library, May 1998
A:Description: Arabidopsis thaliana chromosome II BAC T27B13 genomic sequence.
A:Reference number: Z14178
A:Accession: T00591
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-712 <ROU>
C:Cross-references: EMBL:AC004165; NID:g3150396; PID:g3150413
A:Experimental source: cultivar Columbia
R:Kounisley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul

submitted to the EMBL Data Library, August 1998

A:Description: Arabidopsis thaliana chromosome II BAC F23f1 genomic sequence.

A:Reference number: Z14675

A:Accession: T02479

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-712 <STO>

A:Cross-references: EMBL:AC004680; NID:g3420043; PID:g3420045

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, L.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A04420; MUID:20083487; PMID:10617197

A:Accession: B04704

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-712 <STO>

A:Cross-references: GB:AE002093; NID:g3150413; PID:AAC16965.1; GSPDB:GN00139

C:Genetics:

A:Gene: KUP1; At2g30070; T27B13.19; F23F1.2

A:Map position: 2

C:Interons: 15/3; 97/1; 179/1; 196/1; 283/1; 300/3; 339/1; 424/1

C:Superfamily: barley probable potassium transport protein HAK1

Query Match

Best Local Similarity 54.8%; Score 40; DB 2; Length 712;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 3 EMYNFRSQSLASV 15

|||:|:|:|:|

Db 607 EMYDFSRSLVSAI 619

RESULT 9

F97717

hypothetical protein RC0142 [imported] - Rickettsia conorii (strain Malish 7)

C:Species: Rickettsia conorii

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001

A:Accession: F97717

R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R

Science 293, 2093-2098, 2001

A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.

A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: F97717

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-993 <KUP>

A:Cross-references: GB:AE006914; PIDN:AAL02680.1; PID:G15619186; GSPDB:GN00173

C:Genetics:

A:Gene: RC0142

Query Match

Best Local Similarity 54.8%; Score 40; DB 2; Length 993;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 MYNFRSOLA 12

|||||

Db 716 MYNFSOLA 724

RESULT 10

F71719

hypothetical protein RP104 - Rickettsia prowazekii

C:Species: Rickettsia prowazekii

C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000

R:Anderson, S.G.E.; Zomorodipour, A.; Anderson, J.O.; Slicheritz-Ponten, T.; Almark, U

Nature 396, 133-140, 1998

A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A:Reference number: A11630; MUID:39039499; PMID:9623893

A:Accession: F71719

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-1124 <AND>
A:Cross-references: GB:AJ235270; GB:AJ235269; NID:g3860572; PIDN:CA14573.1; PID:e134241
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: RP104

Query Match 54.8%; Score 40; DB 2; Length 1124;
Best Local Similarity 88.9%; Pred. No. 38;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 MYNFRSOLA 12
||| |||
Db 717 MYNFRSOLA 725

RESULT 11

AI1084
Hypothetical protein lmo0079 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AI1084
R:Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Entian, K.D.; Fsihl, H. D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Author: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Me Ok, C.; Schueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.
Article: Comparative genomics of Listeria species.
A:Reference number: ABI077; MUID:21537279; PMID:11679669
A:Accession: AI1084
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-492 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC98294.1; PID:G16409438; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo0079

Query Match 53.4%; Score 39; DB 2; Length 492;
Best Local Similarity 58.3%; Pred. No. 23;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DSEMYNFRSOLA 12
: : ||| |||
Db 127 EGYMYNFRSOLA 138

RESULT 12

S67564
Hypothetical protein YDL031w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein D2770
C:Species: Saccharomyces cerevisiae
C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002
C:Accession: S67564
R:Paulin, L.; Saren, A.M.; Laamanen, P.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67560
A:Accession: S67564
A:Molecule type: DNA
A:Residues: 1-995 <PAU>
A:Cross-references: EMBL:Z74079; NID:g1431009; PIDN:CA98590.1; PID:e25286; PID:g143101
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:DBP10
A:Cross-references: SGD:S0002189
A:Map position: 4L
C:Keywords: ATP; nucleotide binding; P-loop
F:181-188/Region: nucleotide-binding motif A (P-loop)
F:284-289/Region: nucleotide-binding motif B
F:288-291/Region: DEAD motif

Query Match 53.4%; Score 39; DB 2; Length 995;

Best Local Similarity 46.2%; Pred. No. 51;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
Qy 3 EMYNFRSOLASV 15
::|||::|||:
Db 474 QLYNFRAGLSTL 486

RESULT 13

S57380
Probable membrane protein YOL089c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein O0938
C:Species: Saccharomyces cerevisiae
C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 19-Apr-2002
C:Accession: S57380; S66783; S50416
R:Zumstein, E.; Pearson, B.M.; Kalogeropoulos, A.; Schweizer, M.
Yeast 11, 975-986, 1995
A:Title: A 29.425 kb segment on the left arm of yeast chromosome XV contains more than t
A:Reference number: S57374; MUID:96021609; PMID:8533473
A:Accession: S57380
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1030 <ZUM>
A:Cross-references: EMBL:X83121; NID:9600461; PIDN:CA58190.1; PID:9600469
R:Zumstein, E.; Pearson, B.M.; Kalogeropoulos, A.; Schweizer, M.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S66775
A:Accession: S66783
A:Molecule type: DNA
A:Residues: 1-1030 <ZUM>
A:Cross-references: EMBL:Z74631; NID:g1419931; PIDN:CA99101.1; PID:e251886; PID:g141993
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:HAL9
A:Cross-references: SGD:S0005449
A:Map position: 15L
C:Superfamily: unassigned GAL4-type zinc cluster proteins; GAL4 zinc binuclear cluster h
C:Keywords: transmembrane protein
F:131-171/Domain: GAL4 zinc binuclear cluster homology <GAL4>
F:762-778/Domain: transmembrane #status predicted <TM>

Query Match 53.4%; Score 39; DB 2; Length 1030;
Best Local Similarity 53.3%; Pred. No. 53;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DSEMYNFRSOLASV 15
||| ||| ||| |||
Db 954 DSYLYPIRSDLTSLV 968

RESULT 14

H64752
Probable transcription regulator yagI - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: H64752
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: H64752
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-252 <BLAT>
A:Cross-references: GB:AE000135; GB:U00096; NID:g1786465; PIDN:AACT73375.1; PID:g1786468,
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: yagI
C:Superfamily: acetate operon repressor
C:Keywords: DNA binding; transcription regulation
F:186-207/Region: helix-turn-helix motif

Query Match 52.1%; Score 38; DB 2; Length 252;
 Best Local Similarity 63.6%; Pred. No. 16;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DSEMYNFRSQQL 11
 ||:|||||
 DB 227 DAELANFRQQL 237

RESULT 15

T16257

Hypothetical protein F35C8.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T16257

R:Hu, X.

Submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid F35C8.

A:Reference number: Z18486

A:Accession: T16257

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-312 <WUX>

A:Cross-references: EMBL:U40941; NID:G1072184; PID:G1072190; PIDN:AAA81712.1; CESP:F35C8

C:Genetics:

A:Gene: CESP:F35C8.8

A:introns: 18/31 40/2; 72/2; 113/2; 236/1; 269/1

Query Match

Best Local Similarity 52.1%; Score 38; DB 2; Length 312;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 EMYNFRSQQLASVY 15
 ||:|||||:
 DB 39 EMYNFRSQQLQIV 51

Search completed: November 26, 2003, 12:36:07
 Job time : 14.9157 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:25:10 ; Search time 33.7108 Seconds
(without alignments)
118.334 Million cell updates/sec

Title: US-09-230-111c-16

Perfect score: 73

Sequence: 1 DSEMYNFRSQLASV 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP viirus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	68.5	216	3	014042
2	41	56.2	109	4	081UB7
3	41	56.2	109	4	081UB6
4	41	56.2	804	4	09UTZ6
5	41	56.2	805	4	09BIF1
6	41	56.2	805	4	09NKA7
7	40	54.8	193	5	09W5S5
8	40	54.8	356	17	097Y08
9	40	54.8	712	10	022397
10	40	54.8	712	10	050043
11	40	54.8	993	16	092UC5
12	40	54.8	1124	16	092E44
13	39	53.4	33	16	08EGQ4
14	39	53.4	321	16	08EGQ4
15	39	53.4	492	16	08YAN8
16	39	53.4	840	15	09WBT0

17	39	53.4	1030	3	012180	012180 saccharomyc
18	38	52.1	312	5	020029	020029 caenorhabdi
19	38	52.1	331	6	09TSM4	09TSM4 macaca faec
20	38	52.1	510	16	08CXK7	08CXK7 oceanobacti
21	38	52.1	512	8	09B8G9	09B8G9 heterodoxus
22	38	52.1	521	11	08BSJ2	08BSJ2 mus musculi
23	38	52.1	747	5	08IKO8	08IKO8 plasmodium
24	37	50.7	229	12	010689	010689 cymbidium m
25	37	50.7	233	12	09Q3U7	09Q3U7 cymbidium m
26	37	50.7	233	12	037154	037154 cymbidium m
27	37	50.7	257	17	0973X9	0973X9 sulfolobus
28	37	50.7	342	2	08KJ83	08KJ83 rhizobium 1
29	37	50.7	397	2	030857	030857 vibrio chol
30	37	50.7	649	10	09ME6E	09ME6E nicotiana t
31	37	50.7	938	9	09TIE3	09TIE3 lactobacilli
32	37	50.7	1249	16	08DT19	08DT19 streptococc
33	37	50.7	1326	5	08IB98	08IB98 plasmodium
34	37	50.7	1667	5	018901	018901 caenorhabdi
35	37	50.7	4533	5	09BIX3	09BIX3 tetrahymena
36	36.5	50.0	4589	5	076506	076506 tetrahymena
37	36	49.3	183	6	08MK94	08MK94 echinops te
38	36	49.3	234	16	097KU4	097KU4 clostridium
39	36	49.3	239	5	096618	096618 caenorhabdi
40	36	49.3	263	16	08Z491	08Z491 salmoneilla
41	36	49.3	282	16	092J18	092J18 rickettsia
42	36	49.3	301	5	08IS88	08IS88 leishmania
43	36	49.3	371	2	08KZ35	08KZ35 uncultured
44	36	49.3	376	17	097WK1	097WK1 sulfolobus
45	36	49.3	388	16	09A0P6	09A0P6 streptococc

ALIGNMENTS

RESULT 1

014042

PRELIMINARY;

PRT; 216 AA.

AC 014042;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical conserved protein.
GN SPAC2C4.09
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Barrett B.G., Rajandream M.A., Wood V.;
RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z99259; CAB16369.1; -
DR GeneDB; SPombe; SPAC2C4.09; -
KW Hypothetical protein.
SQ SEQUENCE 216 AA; 25318 MW; B47C549BA8BD420 CRC64;

Query Match

Best local similarity 68.5%; Score 50; DB 3; Length 216;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSEMYNFRSQLASV 13
DB 176 DSEMYNFRSQLASV 188

RESULT 2

ID	081UB7	PRELIMINARY;	PRT;	109	AA.
AC	081UB7.				
DT	01-MAR-2003 (TREMBLrel. 23, Created)				
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)				
DE	01-MAR-2003 (TREMBLrel. 23, Last annotation update)				
DE	CD95 antigen (Fragment).				
GN	CD95.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Euteria; Primates; Carnivora; Hominoidea; Homo.				
OX	NCBI_TaxId=9606;				
RP	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEBLIN=22404279; PubMed=12516573;				
RA	Kurtz J., Perlick A., Schmitz R., Iking-Konert C., Chiorazzi N.,				
RA	Thompson K.M., Winkler T., Ravejsky R., Kueppers R.;				
RT	"Lack of deleterious somatic mutations in the CD95 gene of				
RT	plasmablasts from systemic lupus erythematosus patients and				
RL	autoantibody-producing cell lines.";				
RT	Eur. J. Immunol. 32:3785-3792(2002).				
FT	EMBL, AJ509181; CAD48931.1; --.				
FT	NON TER				
SO	SEQUENCE				
SO	109 AA; 12372 MW; 5EAD927A23C0AAB8 CRC64;				
Query Match					
Best Local Similarity	56.2%;	Score 41;	DB 4;	Length 109;	
Matches	8; Conservative	53.3%;	Pred. No. 6.5;		
		4; Mismatches	3; Indels	0; Gaps	
Cy	1 DSEMYNFRSQASVY 15				
	::: :				
DB	95 DSENNFRNFIQSLV 109				

```

RESULT 3
081UB6
ID 081UB6 PRELIMINARY; PRT; 109 AA.
AC 081UB6;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE CD95 antigen (Fragment).
GN CD95.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RK MEDLINE=22404279; PubMed=12516573;
RA Kuth U., Pernick A., Schmitz R., Iking-Konert C., Chiorazzi N.,
RA Thompson K.M., Winkler T., Rajewsky K., Kuipers R.;
RT "Lack of deleterious somatic mutations in the CD95 gene of
RT plasmablasts from systemic lupus erythematosus patients and
RT autoantibody-producing cell lines";
RL Eur. J. Immunol. 32:3785-3792(2002).
DR EMBL; AJ509182; CAD48932.1; -.
FT NON TER 1
SQ SEQUENCE 109 AA; 12383 MW; D48D9C942D2044B6 CRC64;

Query Match 56.2%; Score 41; DB 4; Length 109;
Best local Similarity 53.3%; Pred. No. 6.5;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0.

Qy 1 DSEMYNFRSOLASV 15
||| |||::|:|
Db 95 DSENSNFRNLTQSLV 109

RESULT 4
Q9UPZ6 PRELIMINARY; PRT; 804 AA.
ID Q9UPZ6
AC Q9UPZ6;

```

```

DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
GN DPF2PA34A014.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RL Submitted R, Haubner D., Mewes H.W., Gassenhuber J., Wiemann S.,
EMBL; AL110224; CAB53682.1; -.
DR MEROPS; M02.006; -.
DR InterPro; IPR001548; Peptidase_M2.
DR InterPro; IPR006025; Zn_M1peptidase.
DR Pfam; PF01401; Peptidase_M2; 1.
DR PRINTS; PR00791; PEBDIP1ASEA.
DR ProDom; PD004184; Peptidase_M2; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 804 AA; 92340 MW; 91FF391074C85DA9 CRC64;

Query Match 56.2%; Score 41; DB 4; Length 804;
Best Local Similarity 66.7%; Pred.No. 53;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 DSEMYNFRSQA 12
|:|:|:|:|:|:|:|
DB 636 DSEMYLFRSSVA 647

```

RESULT 5			
ID	Q9BYF1	PRELIMINARY;	PRT; 805 AA.
AC	Q9BYF1;		
DT	01-JUN-2001	(TREMBLrel. 17, Created)	
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)	
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)	
DE	ACE2.		
GN	ACE2.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxId=9606;		
PN	[1]		
RP	SEQUENCE FROM N.A.		
RP	Komatsu T., Suzuki Y., Sugano S.;		
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Suzuki Y., Matanabe M., Sugano S.;		
RT	"Cloning, expression analysis and chromosomal localization of a novel		
RL	ACE like enzyme.";		
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB046569; BAB40370.1; -.		
DR	MEROPS; M02.006; -.		
DR	Genew; HGNC:13557; ACE2.		
DR	InterPro; IPR001548; Peptidase_M2.		
DR	InterPro; IPR006025; Zn_MTpeptidase.		
DR	Pfam; PF01401; Peptidase_M2; 1.		
DR	PRINTS; PD00791; PEPTIDTASEA.		
DR	ProDom; PD004184; Peptidase_M2; 1.		
DR	PROSITE; PS00142; ZINC_PROTEASE; 1.		
SQ	SEQUENCE 805 AA; 92491 MW; D2AAB4C27088EB72 CRC64;		
Query Match 56.2%; Score 41; DB 4; Length 805;			
Best Match Similarity 66.7%; Pred. No. 53;			
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0			

OY 1 DSEMYNFRSOLA 12
 Db 637 DNEWYLFRRSSVA 648

RESULT 6
 Q9NRA7 PRELIMINARY; PRT; 805 AA.

AC Q9NRA7; 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Angiotensin converting enzyme-like protein (ACE)-related
 DE carboxypeptidase (ACE2).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymphoma;
 RA Turner S.R., Hooper N.M., Hyde R.J., Christie G., Karren E.,
 RA Turner A.J.;
 RT "A Human Homolog of Angiotensin Converting Enzyme - Cloning and
 RT Functional Expression As A Captopril-insensitive Carboxypeptidase";
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RA Donoghue M., Hsieh F., Baronas E., Godbout K., Gosselin M.,
 RA Stagliano N., Donovan M., Woolf B., Robison K., Jeyaseelan R.,
 RA Breitbart R.E., Acton S.;
 RT "A novel ACE-related carboxypeptidase (ACE2) converts angiotensin I to
 RT angiotensin-1-9";
 RT Cite. Res. 0:0-0(2000).
 DR EMBL; AF241254; AAF78220.1; -;
 DR EMBL; AF291820; AAF93721.1; -;
 DR MEROPS; M02.006; -;
 DR InterPro; IPR001548; Peptidase_M2.
 DR InterPro; IPR006025; Zn_MTPeptide.
 DR Pfam; PF01401; Peptidase_M2_1.
 DR PRINTS; PR00791; PEPDIPASEA.
 DR ProDom; PD004184; PEPDIPASEA_M2_1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Carboxypeptidase.
 SQ SEQUENCE 805 AA; 92462 MW; 8E6E80A931550E8 CRC64;

Query Match 56.2%; Score 41; DB 4; Length 805;
 Best Local Similarity 66.7%; Pred. No. 53;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 DSEMYNFRSOLA 12
 Db 637 DNEWYLFRRSSVA 648

RESULT 7
 Q9W5S5 PRELIMINARY; PRT; 193 AA.

AC Q9W5S5; 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE CG17506 protein.
 GN CG17506.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Gelinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zheng L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu B., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Daynoff L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downe M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertilla S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paetle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 DR EMBL; AB002828; AAF45451.1; -;
 DR FLYbase; FBgn040041; CG17506.
 SQ SEQUENCE 193 AA; 22534 MW; FB8201EEFFCD40B8 CRC64;

Query Match 54.8%; Score 40; DB 5; Length 193;
 Best Local Similarity 46.7%; Pred. No. 18;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 DSEMYNFRSOLA 15
 Db 170 DNEWYLFRRSSVA 184

RESULT 8
 Q97Y08 PRELIMINARY; PRT; 356 AA.

AC Q97Y08; 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical protein SSO1545.
 GN SSO1545.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RA MEDLINE=21332296; PubMed=11427726;
 RA She Q., Singh R.K., Contaloni F., Zivanovic Y., Allard G.,
 RA Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moers A., Erasus G., Fletcher C., Gordon P.M.K.,
 RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,

```

RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840 (2001).
DR EMBL, AE006769; AAK41763.1; -.
DR InterPro; IPR004309; DUF238.
DR Pfam; PF03075; DUF238; 1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 356 AA; 41712 MW; 1B791F963C670F4E CRC64;

Query Match 54.8%; Score 40; DB 17; Length 356;
Best Local Similarity 63.6%; Pred. No. 35;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSEMYNFRSOL 11
Db 321 DSEIYNVLTQL 331

RESULT 9
O22397 PRELIMINARY; PRT; 712 AA.
AC O22397;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Putative potassium transporter Atk1p.
GN ATK1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98010480; PubMed=9350997;
RA Quintero F.J., Blatt M.R.;
RT "A new family of K+ transporters from Arabidopsis that are conserved
RT across phyla";
RL FEBS Lett. 415:206-211 (1997).
DR EMBL, AF012656; AAC49844.1; -.
DR InterPro; IPR003855; K+-transporter.
DR Pfam; PF02705; K_trans; 1.
DR TIGRFAMs; TIGR00794; Knp; 1.
SQ SEQUENCE 712 AA; 79032 MW; BF766100E4881A7B CRC64;

Query Match 54.8%; Score 40; DB 10; Length 712;
Best Local Similarity 53.8%; Pred. No. 72;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 EMYNFRSOLASV 15
Db 607 EMYDFESRLVSAI 619

RESULT 10
O50043 PRELIMINARY; PRT; 712 AA.
AC O50043;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE High-affinity potassium transporter (high affinity K+ transporter)
DE (AtKUP1 AtK1p).
GN KUP1 OR T27E13.19 OR AT2G30070.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;

```

```

RA Kim E.J., Kwak J.M., Uozumi N., Schroeder J.I.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Fu H.H., Luan S.;
RL Plant Cell 0:0-0 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,
RA Sykes S.M., Kaul S., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF029876; AAB88901.1; -.
DR EMBL, AF031118; AAB87687.1; -.
DR EMBL, AC004165; AAC16965.1; -.
DR EMBL, AC004680; AAM14984.1; -.
DR InterPro; IPR003855; K+-transporter.
DR Pfam; PF02705; K_trans; 1.
DR TIGRFAMs; TIGR00794; Knp; 1.
SQ SEQUENCE 712 AA; 79126 MW; 0FE32BE88807536D CRC64;

```

```

Query Match 54.8%; Score 40; DB 10; Length 712;
Best Local Similarity 53.8%; Pred. No. 72;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 EMYNFRSOLASV 15
Db 607 EMYDFESRLVSAI 619

RESULT 11
O92JC5 PRELIMINARY; PRT; 993 AA.
AC O92JC5;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein RCO142.
GN RCO142.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxId=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audifren P., Fournier P.-E., Barbe V.,
RA Raoult D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098 (2001).
DR EMBL, AE008581; AAL02680.1; -.
DR InterPro; IPR001092; HLM_dasic.
DR PROSITE; PS00038; HLM_1; 1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 993 AA; 108744 MW; 35856A19653A53A5 CRC64;

```

```

Query Match 54.8%; Score 40; DB 16; Length 993;
Best Local Similarity 88.9%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 4 MYNFRSOLA 12
 |||||
 Db 716 MYNFRSOLA 724

RESULT 12

Q9ZE44

ID Q9ZE44 PRELIMINARY; PRT; 1124 AA.

AC Q9ZE44; 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Hypothetical protein RP104.

GN RP104.

OS Rickettsia prowazekii.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;

OC Rickettsiaceae; Rickettsiae; Rickettsia.

OX NCBI_TaxID=782;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Medrid E;

RX MEDLINE=9903949; PubMed=9823893;

RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,

RA Sichteritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,

RA Eriksson A.-S., Winkler H.H., Kurland C.G.;

RT "The genome sequence of Rickettsia prowazekii and the origin of

RL mitochondria."

RL Nature 396:133-140(1998).

DR EMBL; AJ235270; CAI4573.1; -.

DR InterPro; IPR001092; HLH_bact.

DR PROSITE; PS00038; HLH_1.1.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 1124 AA; 124974 MW; EACIE61564734EBB CRC64;

Query Match 54.8%; Score 40; DB 16; Length 1124;
 Best Local Similarity 88.9%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 MYNFRSOLA 12
 |||||
 Db 717 MYNFRSOLA 725

RESULT 13

Q8EGQ4

ID Q8EGQ4 PRELIMINARY; PRT; 33 AA.

AC Q8EGQ4; 01-MAR-2003 (TREMBLrel. 23, Created)

DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Hypothetical protein.

GN SOL1544.

OS Shewanella oneidensis.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;

OC Alteromonadaceae; Shewanella.

OX NCBI_TaxID=70863;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=WR-1;

RX MEDLINE=22397686; PubMed=12368813;

RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,

RA Read T.D., Eisen J.A., Seshadri R., Ward N., Mehe B., Clayton R.A.,

RA Meyer T., Tsaplin A., Scott J., Beaman M., Brinkac L., Daugherty S.,

RA Deboy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,

RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,

RA Vamathevan J., Weidman J., Imbraim M., Lee K., Berry K., Lee C.,

RA Mueller J., Khouri H., Gill J., Uterback T.R., McDonald L.A.,

RA Feldblyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.,

RT "Genome sequence of the dissimilatory metal ion-reducing bacterium

RT Shewanella oneidensis."

RT Nat. Biotechnol. 20:1118-1123(2002).

DR EMBL; AE015600; AAN54602.1; -.

DR TIGR; SO1544; -.

KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 33 AA; 3611 MW; 1F20EFP390BD21C4 CRC64;

Query Match 53.4%; Score 39; DB 16; Length 33;
 Best Local Similarity 54.5%; Pred. No. 4.4;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DSEMYNFRSOLA 11
 ||::|||::|

Db 19 DNDLYNFRSPL 29

RESULT 14

Q8EZD0

ID Q8EZD0 PRELIMINARY; PRT; 321 AA.

AC Q8EZD0; 01-MAR-2003 (TREMBLrel. 23, Created)

DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE NAD(P)-dependent steroid dehydrogenase (EC 1.1.1.-).

GN LA3924.

OS Leptospira interrogans.

OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.

OX NCBI_TaxID=173;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lat;

RX Ren S.;

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE011548; AAN51122.1; -.

DR Oxidoreductase; Complete proteome.

SQ SEQUENCE 321 AA; 35515 MW; 0E4CDB5B7A9D0FC CRC64;

Query Match 53.4%; Score 39; DB 16; Length 321;
 Best Local Similarity 46.7%; Pred. No. 48;
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DSEMYNFRSOLA 15
 ||::|||::|

Db 225 DDEIFNFRFLESL 239

RESULT 15

Q8YAN8

ID Q8YAN8 PRELIMINARY; PRT; 492 AA.

AC Q8YAN8; 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Hypothetical protein lmo0079.

GN LMO0079.

OS Listeria monocytogenes.

OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

OX NCBI_TaxID=1639;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=EGD-e / Serovar 1/2a;

RX MEDLINE=21537279; PubMed=11679659;

RA Glaeser P., Frangoul L., Buchrieser C., Ruenliok C., Amend A.,

RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,

RA Chabir A., Chetouani F., Couve E., de Daruvar A., Deloux P.,

RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,

RA Gauthier L., Geisel W., Garcia-del Portillo F., Garrido P.,

RA Jones L.-M., Kaerst U., Kuhn M., Kunz F., Kurapkat G.,

RA Madueno E., Maitournan A., Mata Vicente J., Ng E., Nedjati H.,

RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,

RA Remmel B., Rose M., Schueter T., Simoes N., Tietz A.,

RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;

RT "Comparative genomics of Listeria species."

RT Science 294:849-852(2001).

DR EMBL; AL591973; CAC98294.1; -.

DR ListLib; LMO00079; -.

KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 492 AA; 55280 MW; 977C6549713B01C7 CRC64;

Query Match 53.4%; Score 39; DB 16; Length 492;
Best Local Similarity 58.3%; Pred. No. 75;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 DSEMYNFRSOLA 12
: : : : :
Db 127 EGMOMYRSOLA 138

Search completed: November 26, 2003, 12:34:41
Job time : 34.7108 secs